

**MULTIPLEX-RT-PCR-SOVELLUSTEN KEHITTÄMINEN KAS-
VIVIRUSDIAGNOSTIIKKA VARTEN JA VIRUSTEN TUNNIS-
TAMINEN SIRNA-ANALYYSILLÄ**

Jaana Jukkala
Maisterintutkielma
Helsingin yliopisto
Maataloustieteiden osasto
Kasvipatologia
Tammikuu 2019

Tiedekunta — Fakultet — Faculty		Laitos — Institution — Department	
Maatalous-metsätieteellinen tiedekunta		Maataloustieteiden osasto	
Tekijä — Författare — Author			
Jaana Jukkala			
Työn nimi — Arbetets titel — Title			
Multiplex-RT-PCR-sovellusten kehittäminen kasvivirusdiagnostiikkaa varten ja virusten tunnistaminen siRNA-analyysillä			
Oppiaine — Läroämne — Subject			
Kasvipatologia			
Työn laji — Arbetets art — Level		Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Maisterintutkielma		Tammikuu 2019	105
Tiivistelmä — Referat — Abstract			
<p>Multiplex-PCR-menetelmässä käytetään samassa reaktiossa useita, eri DNA-jaksoille spesifisiä alukkeita, mikä mahdollistaa esimerkiksi useiden eri virusten tai virusryhmien tunnistamisen samalla PCR-reaktiolla. Uuden laaja-alaisen kasvivirusten tunnistusmenetelmän, siRNA-analyysin, odotetaan edistävän kasvivirusten tunnistamista, sillä se ei vaadi ennakkoletusta näytteessä mahdollisesti esiintyvistä viruksista. Menetelmän ansiosta testattavien virusten rajaukset käyvät tarpeettomiksi.</p> <p>Tämän tutkimuksen tavoitteena oli optimoida Eviran (nyk. Ruokavirasto) kasvintuhoojalaboratoriolle kaksi tai useampi yksivaiheinen multiplex-RT-PCR-testi yhdeksälle, kirjallisuudesta etsitylle, degeneroidulle alukeparille poty-, potex-, tobra- ja tospoviruskujen, tobamoviruksen alaryhmä yhden, nepoviruksen a- ja b-alaryhmien, <i>Bromoviridae</i>-virusheimon ja <i>Tombusviridae</i>-heimoon kuuluvien carmo-, diantho- ja tombusviruskujen samanaikaiseen tunnistamiseen.</p> <p>Tutkimuksessa käytettiin sekä multiplex-RT-PCR-testejä että siRNA-syväsekvensointia selvittämään viroottisista ja viroottisilta vaikuttaneista näytteistä esiintyvät kasvivirukset. Tunnistamalla kasveissa esiintyvät virukset haluttiin arvioida multiplex-RT-PCR-testien soveltuvuutta useiden eri RNA-viruskuihin kuuluvien virusten tunnistamiseen samalla testillä. siRNA-menetelmää käytettiin multiplex-RT-PCR-testien tulosten oikeellisuuden varmistamiseen.</p> <p>Tutkimuksessa pystyttiin luomaan kaksi multiplex-RT-PCR-testiä, jotka tunnistavat yhteensä kahdeksaan eri virusryhmään kuuluvia viruksia. Saadut tulokset toimivat lähtökohtana multiplex-RT-PCR-testien validoinnille, jonka avulla voidaan varmistua testien antamien tulosten luotettavuudesta ja soveltuvuudesta käyttötarkoitukseensa. Tulokset osoittivat siRNA-syväsekvensoinnin tunnistaneen osasta löytyneistä viruksista lähes niiden koko genomin.</p> <p>Kartoitetuista kasvinäytteistä optimoidut multiplex-RT-PCR-testit havaitsivat lupiinin mosaikkiviruksen (LuMV) lupiinista (<i>Lupinus polyphyllus</i> Lindl.), arabiksen mosaikkiviruksen (ArMV) kiinanjaloangervosta (<i>Astilbe chinensis</i> (Maxim.) Franch.) ja kvinoasta (<i>Chenopodium quinoa</i> Willd.), tupakan rattleviruksen (TRV) suikeroalpista (<i>Lysimachia nummularia</i> L.) ja mahdollisesti uuden potyviruslajin valkomesikästä (<i>Melilotus albus</i> Medik.). siRNA-syväsekvensointi tunnistasi näytteistä samat kasvivirukset, mikä vahvistaa luottamusta optimoituja multiplex-RT-PCR-testejä kohtaan.</p>			
Avainsanat — Nyckelord — Keywords			
Multiplex-RT-PCR, siRNA-syväsekvensointi, kasvivirukset, RNA-hiljennys, Velvet, VirusDetect			
Säilytyspaikka — Förvaringsställe — Where deposited			
HELDA – Helsingin yliopiston digitaalinen arkisto			
Muuta tietoja — Övriga uppgifter — Additional information			
Työtä ohjasivat Johanna Santala (Ruokavirasto), Minna Rajamäki (Helsingin yliopisto) ja Jari Valkonen (Helsingin yliopisto)			

Tiedekunta — Fakultet — Faculty Faculty of Agriculture and Forestry		Laitos — Institution — Department Department of Agricultural Sciences	
Tekijä — Författare — Author Jaana Jukkala			
Työn nimi — Arbetets titel — Title Developing Multiplex PCR assays for plant virus diagnostics and detection of viruses using siRNA deep sequencing			
Oppiaine — Läroämne — Subject Plant pathology			
Työn laji — Arbetets art — Level Master's thesis		Aika — Datum — Month and year January 2019	Sivumäärä — Sidoantal — Number of pages 105
<div>Tiivistelmä — Referat — Abstract</div> <p>Multiplex PCR uses several primer pairs, each specific for different DNA sequences in the same reaction, enabling viruses from different families and genera to be recognized by the same test. The new broad-based method of detecting plant viruses, siRNA analysis, is expected to improve the identification of plant viruses as it does not require a prediction of viruses that may be present in the sample. Due to the method, limiting the test to detecting only certain viruses is unnecessary.</p> <p>The aim of this study was to optimize two or more single-phase multiplex RT-PCR tests for the pest laboratory of the Finnish Food Safety Authority (now known as Finnish Food Authority). The tests were for nine published, degenerate primer pairs for the simultaneous identification of <i>Potyvirus</i>, <i>Potexvirus</i>, <i>Tobravirus</i> and <i>Tospovirus</i> genera, <i>Tobamovirus</i> subgroup 1, <i>Nepovirus</i> subgroups a and b, <i>Bromoviridae</i> family and <i>Carmovirus</i>, <i>Dianthovirus</i>, and <i>Tombusvirus</i>, which belong in the <i>Tombusviridae</i> family.</p> <p>Both multiplex RT-PCR tests and siRNA deep sequencing were used to detect plant viruses from infected samples or samples showing viral symptoms. By identifying the viruses in the plant samples, the goal was to estimate the suitability of multiplex RT-PCR tests for identifying multiple RNA viruses from different genera. siRNA analysis was used to ensure the correctness of the multiplex RT-PCR results.</p> <p>In this study, two multiplex RT-PCR tests were created to detect viruses belonging to eight different virus groups. The obtained results serve as a basis for the validation of the multiplex RT-PCR tests. The validation is required to verify the suitability of the method for its intended use and the reliability of the results of the method. The results showed that siRNA deep sequencing was able to detect almost the whole genome from some of the found viruses.</p> <p>Multiplex RT-PCR tests detected <i>Lupine mosaic virus</i> (LuMV) from lupine (<i>Lupinus polyphyllus</i> Lindl.), <i>Arabic mosaic virus</i> (ArMV) from Chinese astilbe (<i>Astilbe chinensis</i> (Maxim.) Franch.) and from quinoa (<i>Chenopodium quinoa</i> Willd.), <i>Tobacco rattle virus</i> (TRV) from moneywort (<i>Lysimachia nummularia</i> L.) and possibly a new <i>Potyvirus</i> species from honey clover (<i>Melilotus albus</i> Medik.). siRNA analysis detected the same viruses, which increases confidence in the optimized multiplex RT-PCR tests.</p>			
<div>Avainsanat — Nyckelord — Keywords</div> <p>Multiplex-RT-PCR, siRNA deep-sequencing, plant viruses, RNA silencing, Velvet, VirusDetect</p>			
<div>Säilytyspaikka — Förvaringsställe — Where deposited</div> <p>HELDA – Digital Repository of the University of Helsinki</p>			
<div>Muita tietoja — Övriga uppgifter — Additional information</div> <p>Supervised by Johanna Santala (Finnish Food Authority), Minna Rajamäki (University of Helsinki) and Jari Valkonen (University of Helsinki)</p>			

Sisältö

1	Lyhenteet ja käsitteet	6
2	Johdanto	8
2.1	Kasvivirukset	9
2.1.1	Kasvivirusten monistuminen ja liikkuminen kasvissa	9
2.1.2	Kasvivirusten leviäminen ja virusten aiheuttamat oireet . .	10
2.1.3	Kasvivirusten luokittelu	11
2.2	Kasvivirusten tunnistusmenetelmät	11
2.2.1	Indikaattorikasvien käyttö kasvivirusten tunnistamisessa . .	13
2.2.2	Serologiset tunnistusmenetelmät	14
2.2.3	Polymeraasiketjureaktio (PCR)	14
2.2.3.1	PCR-menetelmän suunnittelu ja optimointi	16
2.2.3.2	Degeneroitujen alukkeiden käyttö kasvivirusdiag- nostiikassa	17
2.2.3.3	Multiplex-PCR:n käyttö kasvivirusdiagnostiikassa	18
2.2.4	Syväseksensoinnin käyttö kasvivirusdiagnostiikassa	20
2.2.5	siRNA-syväseksensoinnin käyttö kasvivirusdiagnostiikassa	22
2.2.5.1	Virustartunnan aktivoima RNA-hiljennys	24
2.2.5.2	Virusten puolustautuminen RNA-hiljennystä vastaan	26
2.3	Kasvivirusten esiintyminen luonnonvaraisissa kasveissa	27
2.4	Kasvinterveyden ja taimiaineiston valvonta	28
3	Tutkimuksen tavoitteet	29
4	Tutkimusaineisto ja -menetelmät	30
4.1	Kasvimateriaali	30
4.2	RNA:n eristäminen multiplex-RT-PCR-testiä varten	32
4.3	Multiplex-RT-PCR:n optimointi ja suorittaminen	33
4.4	RNA:n eristäminen siRNA-syväseksensointia varten	38
4.5	siRNA-syväseksensointi	39
4.6	Kontigien kokoaminen sRNA-sekvensseistä Velvet-ohjelmalla . . .	40
4.7	Kontigien kokoaminen sRNA-sekvensseistä VirusDetect-ohjelmalla	42
4.8	sRNA-sekvenssien linjaaminen referenssisekvensseihin Bowtie 2 -ohjelmalla	42
4.9	Sekvenssien fylogeneettinen analyysi	43
5	Tulokset	44
5.1	Multiplex-RT-PCR:n optimoinnin tulokset	44
5.2	Multiplex-RT-PCR:llä tunnistetut kasvivirukset	49
5.3	siRNA-syväseksensoinnin tulokset	54
5.4	Velvet-ohjelmalla tunnistetut virukset	54
5.5	VirusDetect-ohjelmalla tunnistetut virukset	57
5.6	Bowtie 2 -ohjelmalla referenssiviruksiin linjatut sRNA-sekvenssit .	62
5.7	Sekvenssien fylogeneettinen analyysi	68

5.8	Yhteenveto multiplex-RT-PCR:llä, Velvet- ja VirusDetect-ohjelmilla tunnistetuista viruksista	74
6	Tulosten tarkastelu	75
7	Johtopäätökset	81
8	Kiitokset	82
	Lähteet	83
	Liitteet	96
	Liite A Vaihtelevat emäket	97
	Liite B Multiplex-RT-PCR-sekvenssit	98
	Liite C Fylogeneettisten puiden referenssisekvenssit	103

1 Lyhenteet ja käsitteet

AGO	ARGONAUTE, yksijuosteista RNA:ta hajottava endoribonukleaasi
ApMV	<i>Apple mosaic virus</i> , omenan mosaiikkivirus
ArMV	<i>Arabidopsis mosaic virus</i> , arabiksen mosaiikkivirus
BLAST	Basic Local Alignment Search Tool
BSMV	<i>Barley stripe mosaic virus</i> , ohran juovamosaiikkivirus
BYMV	<i>Bean yellow mosaic virus</i> , pavun keltamosaiikkivirus
BYV	<i>Beet yellows virus</i> , juurikkaan keltavirus
CI	Cylindrical inclusion, CI-proteiini
CIYVV	<i>Clover yellow vein virus</i> , apilan keltasuonivirus
CP	Capsid protein, kuoriproteiini
DCL	DICER, kaksijuosteista RNA:ta hajottava endoribonukleaasi
DNA	Dioxyribonucleic acid, deoksiribonukleiinihappo
dNTP	Deoxynucleoside triphosphate, deoksiribonukleotiditriposfaatti
ELISA	Enzyme-linked immunosorbent assay, entsyymivälitteinen immuno-sorbenttimääritys
GFLV	<i>Grapevine fanleaf virus</i> , viiniköynnöksen viuhkalehtivirus
HC-Pro	Helper component-proteinase, HC-pro-proteiini
HTS	High throughput sequencing, syväsekvensointi
INSV	<i>Impatiens necrotic spot virus</i> , palsamin kuoliolaikkuvirus
k-mer	tämän tutkielman yhteydessä joukko Velvet-ohjelman katkomia, <i>k</i> -mittaisia nukleotidisekvenssejä
kDa	kiloDalton
LuMV	<i>Lupine mosaic virus</i> , lupiinin mosaiikkivirus
MP	Movement protein, liikkumisproteiini
MV1	<i>Mint virus 1</i> , minttuvirus 1
NCBI	National Center of Biotechnology Information
NIa	Nuclear inclusion a-proteinase, NIa-proteiini

NIb	Nuclear inclusion b-proteinase, NIb-proteiini
PCR	Polymerase chain reaction, polymeerasiketjureaktio
PepMV	<i>Pepino mosaic virus</i> , pepinon mosaiikkivirus
PPV	<i>Plum pox virus</i> , luumun rokkovirus
RdRp	RNA-dependent RNA polymerase, RNA:sta riippuvainen RNA-polymeraasi
RISC	RNA-induced silencing complex, RNA:n indusoima hiljennyskompleksi
RNA	Ribonucleic acid, ribonukleiinihappo
SiRNA	Short interfering RNA, pikku-RNA-molekyyli
SMV	<i>Soybean mosaic virus</i> , soijapavun mosaiikkivirus
sRNA	Small RNA, pikku-RNA; tämän tutkielman yhteydessä 21–24 nukleotidin pituisia sekvenssejä
TBRV	<i>Tomato black ring virus</i> , tomaatin mustarengaslaikkuvirus
TBSV	<i>Tomato bushy stunt virus</i> , tomaatin kitukasvuvirus
TMV	<i>Tobacco mosaic virus</i> , tupakan mosaiikkivirus
TRV	<i>Tobacco rattle virus</i> , tupakan rattlevirus
UTR	Untranslated region, transloitumaton alue
VPg	Viral protein genome linked, VPg-proteiini

2 Johdanto

Kasvivirukset ovat obligaatteja solunsisäisiä parasiitteja, jotka aiheuttavat kasvintuotannolle tuotannollisia tappioita ja torjunta- ja riskinhallintakustannuksia (Langham 2007, Savary ym. 2012). YK:n elintarvike- ja maatalousjärjestö FAO:n esittämän arvion mukaan kasvitautien ja tuholaisten aiheuttamat satotappiot ovat maailmanlaajuisesti 20–40 % vuotuisesta tuotannosta (FAO 2017b, s. 58). Suomessa taloudellisesti haitallisina kasviviruksina pidetään erityisesti perunan Y-virusta (PVY), perunan maltokaarivirusta (PMTV) ja ohran kääpiökasvuvirusta (BYDV) (Tapio ym. 1997).

Kansainvälinen kasvikauppa sekä muutokset ilmasto-olosuhteissa lisäävät riskiä uusien kasvivirusten leviämisestä Suomeen (Lemmetty ym. 2011, Luck ym. 2011). Tästä esimerkkejä ovat vuosien 1997–2010 välisenä aikana Suomessa tehdyt havainnot kahdeksasta uudesta viruslajista puutarhakasveissa ja vuosina 2008–2009 havainnot viidestä uudesta viroidista koristekasveissa ja kasvihuonevihanneksissa (Lemmetty ym. 2011).

Kasvivirusten valvontaa vaikeuttavat niiden suuri määrä ja monimuotoisuus, mistä johtuen viruksia tunnistavia vasta-aineita ja alukkeita on kehitetty tyypillisesti vain taloudellisesti merkittävien viruslajien varalta. Valmiutta kaikkien virusten diagnostiikkaan ei ole mahdollistaa ylläpitää, vaan testauksissa etsittävät tuhoojat on valittava etukäteen (Jones ym. 2017).

Kasvivirusten tunnistamiseen liittyviä haasteita on pyritty helpottamaan multiplex-PCR-menetelmillä, jotka mahdollistavat useiden eri heimoihin ja sukuihin kuuluvien virusten tunnistamisen samalla testillä (Boonham ym. 2014). Uuden laaja-alaisen kasvivirusten tunnistusmenetelmän, siRNA-analyysin, odotetaan edistävän kasvivirusten tunnistamista, sillä se ei vaadi ennakko-oletusta näytteessä mahdollisesti esiintyvistä viruksista. Menetelmän ansiosta testattavien virusten rajaukset käyvät tarpeettomiksi (Kreuze ym. 2009).

Tässä tutkimuksessa tavoitteena oli optimoida Eviran (nyk. Ruokavirasto) kasvintuhoojalaboratoriolle multiplex-RT-PCR-testi yhdeksälle alukeparille poty-, potex-, tobra- ja tospovirussukujen, tobamovirus alaryhmä yhden, nepoviruksen a- ja b-alaryhmien, *Bromoviridae*-virusheimon ja *Tombusviridae*-heimoön kuuluvien carmo-, diantho- ja tombus-virussukujen tunnistamiseen. Tutkimuksessa käytettiin sekä multiplex-RT-PCR-testiä että siRNA-menetelmää selvittämään saaduista viroottisista ja viroottisilta vaikuttaneista näytteistä esiintyvät kasvivirukset.

Tunnistamalla kasveissa esiintyvät virukset haluttiin osaltaan arvioida multiplex-RT-PCR:n soveltuvuutta kasvintuhojalaboratorion käyttöön. siRNA-menetelmää käytettiin multiplex-PCR-testien tulosten oikeellisuuden varmistamisen. Lisäksi saatiin tietoa siRNA-menetelmän käytettävyydestä virusten testaamisessa.

2.1 Kasvivirukset

Kasvivirologiatieteen kehittäjänä pidetään Martinus Willem Beijerinckia, joka vuonna 1897 osoitti tekemässään tutkimuksessa tupakan mosaiikkitaudin johtuvan bakteeria pienemmästä patogeenista (Hull 2014, s. 3). Fraenkel-Conrat (1956) osoitti, että viruspartikkelista puhdistetun RNA-molekyylin siirtyminen isäntäsoluun riittää aiheuttamaan virusinfektion. Tästä voitiin päätellä, että viruksen RNA sisälsi infektiivisyyteen tarvittavan geneettisen informaation. Tupakan mosaiikkiviruksen genomin emäsjärjestys saatiin selvitettyä vuonna 1982 ensimmäisenä kasviviruksena (Goelet ym. 1982). Monia kasviviruksia on käytetty malliesimerkkeinä antamaan tietoa virusten geenien ilmentymisestä ja genomin replikaatiosta. Kasvivirologiatieteessä tehdyt tutkimuslöydökset ovat edesauttaneet genetiikan, genomiikan ja molekyylibiologiatieteen tutkimusta (Lomonossoff 2018).

Kasvivirukset ovat obligaatteja solunsisäisiä parasiitteja, joilla ei esiinny omaa aineenvaihduntaa eikä proteiinien tuottamiseen tarvittavia ribosomeja. Kasviviruspartikkeli koostuu vähintään genomista ja tätä suojaavasta proteiinikuoresta. Kasvivirukset eivät pysty lisääntymään ilman isäntäsolua. Virusten on selviydyttävä solun ulkoisista olosuhteista ja pystyttävä purkautumaan sopivassa isäntäsolussa. Tunkeuduttuaan soluun virukset käyttävät tämän metabolista ja synteettistä laitteistoa monistuaakseen yleensä solun hyvinvoinnin kustannuksella (Langham 2007). Kaikki kasvivirukset valmistavat proteiineja, joita tarvitaan nukleiinihappojen ja proteiinikuoren rakentamiseen sekä liikkumiseen kasvin sisällä (Agrios 2005, s. 149).

2.1.1 Kasvivirusten monistuminen ja liikkuminen kasvilla

Kasvivirusten monistuminen voidaan jakaa eri vaiheisiin, jotka ovat 1) viruksen tarttuminen isäntäsolun pinnalle, 2) viruksen siirtyminen isäntäsoluun, 3) genomin vapautuminen ja replikaatio, 4) proteiinisynteesi, 5) viruksen kokoaminen ja 6) vapautuminen muihin kasvisoluihin ja kasvinosiin (Hull 2014, s. 168–169). Viruksen kopiointitapa riippuu viruksen genomin koostumuksesta. Kasvivirusten genomi koostuu joko yksi- tai kaksijuosteisesta DNA:sta tai yksi- tai kaksijuosteisesta RNA:sta. Yksijuosteinen RNA voi olla positiivis- tai negatiivissäikeistä. Positiivinen

RNA toimii lähetti-RNA:n tavoin, jolloin se kykenee suoraan isäntäsoluun päästyään aloittamaan proteiinisynteesin. Negatiivisten RNA-virusten RNA-polymeraasi valmistaa ensin plusmerkkisiä kopioita miinusmerkkisestä RNA:sta, minkä jälkeen proteiinisynteesi voi alkaa (Agrios 2005, s. 732–733). Suurin osa kasvivirusista on positiivisia yksijuosteisia RNA-virusia (Sastry 2013, s. 16).

Kasvin sisällä virukset kulkevat lyhyet matkat solusta toiseen kasvin soluseinässä olevien aukkojen, plasmodesmojen, kautta. Liikkumista varten kasvivirus koodaavat erityistä liikkumisproteiinia, jotka muokkaavat plasmodesmiä niin, että virukset pääsevät kulkemaan sen läpi. Kasvirusien systeeminen leviäminen eri kasvinosien välillä tapahtuu passiivisesti johtosolukon välityksellä (Hull 2014, s. 531).

2.1.2 Kasvirusien leviäminen ja virusten aiheuttamat oireet

Kasviruset eivät pysty itsenäisesti liikkumaan isännän ulkopuolella tai läpäisemään kasvin soluseinää. Mekaanisessa siirrostamisessa kasvivirus leviävät kasvista toiseen virusia sisältävän kasvimehun välityksellä. Kasviruset leviävät luonnossa viroottisten siementen, siitepölyn tai kantajien (vektorien) välityksellä, joita voivat esimerkiksi olla hyönteiset, sukkulamadot, punkit ja sienimäiset eliöt (Agrios 2005, s. 737–743). Kasviruset pääsevät siirtymään isäntäsoluun avonaisten haavojen kautta tai vektorien rikkoessa kasvin soluseinän rakenteen. Kasviruset eroavat toisistaan erilaisen pysyvyyden ja vektorin valikoivuuden mukaan. Pysymättömät virukset säilyvät vektorissa vain lyhyen aikaa, kun taas pysyvät virukset säilyvät vektorissa koko sen eliniän. Jotkin kasvivirusista leviävät useamman vektorin välityksellä, kun taas joidenkin virusten leviäminen voi olla riippuvaista esimerkiksi jostain tietystä hyönteislajista (Whitfield ym. 2015).

Kasvirusien aiheuttamat oireet riippuvat viruksen, isännän ja ympäristön välisestä vuorovaikutuksesta vaihdellen oireettomista kasveista virusinfektion takia kuoleviin kasveihin. Virustaudin oireet ilmenevät kasvien kitukasvuisuutena, lakastumisena, erilaisina kehityshäiriöinä ja lehtien kierteisyytenä, kirjavuutena tai kurttuisuutena, rengasmaisina kuvioina tai kloroottisina tai nekroottisina laikkuina. Virusinfektio voi olla latentti, jolloin kasvi on ulkoisesti terveen näköinen (Hull 2014, s. 145–153).

2.1.3 Kasvivirusten luokittelu

Kasvivirusten virallisesta luokituksesta vastaa Kansainvälinen virusluokituskomitea (International Committee on Taxonomy of Viruses, ICTV). Kasvivirusten luokittelu perustuu viruspartikkelin rakenteeseen ja sen fyysisiin ja kemiallisiin ominaisuuksiin, nukleiinihappojen rakenteeseen, virusproteiineihin, serologiaan, biologisiin ominaisuuksiin, kuten viruksen isäntäkirjoon ja virusinfektion aiheuttamiin oireisiin, sekä leviämistapaan kasvista toiseen. Kasvivirukset ovat yleensä nimetty ensimmäisenä löydetyn isäntäkasvin ja kasvissa havaittujen oireiden perusteella. Viruslajien määrittelyn hankaluus sekä uuden tiedon kertyminen virusgenomeista vaikeuttavat virusten luokittelua, mistä johtuen osaa viruksista ei ole pystytty vielä sijoittamaan mihinkään virusheimoon ICTV:n järjestelmässä (Hull 2014, s. 22–30).

Viroideja ja satellittiviruksia ei luokitella viruksiin kuuluviksi, mutta ominaisuuksiltaan ne ovat läheisesti virusten kaltaisia. Viroidit ovat muutamasta sadasta nukleotidista koostuvia, rengasmaisia yksijuosteisia RNA-molekyylejä, joilla on sauvamainen sekundäärirakenne. Toisin kuin viruksilla, viroideilla ei esiinny kapsidia tai lipidivaippaa (Hull 2014, s. 199). DNA:sta tai RNA:sta koostuvat satellittivirukset eivät kykene itsenäiseen replikaatioon, vaan ne käyttävät hyväkseen avustavan viruksen proteiineja. Satellittivirukset voivat vaikuttaa auttajaviruksen määrään ja patogeenisuuteen ja siten kasvissa havaittavien oireiden laatuun (Huymer 2009).

2.2 Kasvivirusten tunnistusmenetelmät

Kasviviruksia ei voida torjua sien- tai bakteeritautien tavoin kemiallisilla torjunta-aineilla, mistä johtuen kasvivirusten torjunta perustuu epäsuoriin, viruksia ennalta ehkäiseviin torjuntamenetelmiin. Ennalta ehkäisevät torjuntamenetelmät tähtäävät virusten leviämisen tai niiden infektiolähteiden vähentämiseen. Näitä ovat esimerkiksi viruksia levittävien hyönteisvektoreiden torjunta ja viruksista vapaan, puhtaan taimimateriaalin käyttö (López ym. 2003).

Kasvivirusten varhainen ja luotettava tunnistaminen ehkäisee virusten leviämistä ja niiden aiheuttamia sadonmenetyksiä. Kasvivirusten tunnistamiseen on kehitetty useita eri menetelmiä, jotka pohjautuvat viruksen biologiseen aktiivisuuteen, genomien emäsjärjestykseen, proteiiniakuoren rakenteeseen tai viruspartikkelin fyysisiin ominaisuuksiin (Ward ym. 2004).

Kasvivirusten tunnistusmenetelmillä saadaan tietoa niiden esiintymisestä ympäristössä, mitä Jonesin (2014) mukaan tarvitaan:

- kasvivirusten eri torjuntamenetelmien tehokkuuden arvioinnissa
- viruksista johtuvien satovahinkojen arvioinnissa
- ydinkasviaineiston puhtauden varmistamisessa
- kasvin taudinkestävyyden jalostamisessa
- kasvitautien leviämisen estämiseksi tehtävässä valvonnassa
- kasvitautien leviämisen ja infektiolähteiden ennustamisessa, mallinnuksessa ja riskinarvioinnissa.

Kasvivirusten tunnistusmenetelmän valintaan vaikuttavat Jonesin ym. (2014) ja Lópezin ym. (2003) seuraavat menetelmän eri ominaisuudet:

- nopeus
- kustannukset
- suorittamiseen vaadittava osaaminen
- suorittamiseen vaadittava tila
- soveltuvuus suurien näytemäärien käsittelyyn.

Uudet kasvitautien tunnistusmenetelmät on validoitava ennen niiden käyttöönottoa lakisääteisten näytteiden analysoinnissa. Euroopan unionissa testien validointi tehdään Eurooppalaisen ja Välimerenmaiden kasvinsuojelujärjestön EPPO:n (European and Mediterranean Plant Protection Organization) ohjeiden mukaan (EPPO 2014). Testien validoinnissa huomioitavat ominaisuudet vaikuttavat edellisten esimerkkien ohella tunnistusmenetelmän valintaan. EPPO:n ohjeiden mukaan testien validoinnissa huomioitavia ominaisuuksia ovat:

- herkkyys
 Diagnostinen herkkyys kuvaa positiiviseksi todettujen osuutta odotetuista positiivista näytteistä ja analyttinen herkkyys kuvaa paljonko virusta tarvitaan, jotta se on menetelmällä tunnistettavissa.
- valikoivuus
 Diagnostinen valikoivuus on oikeiden negatiivisten tulosten suhde oikeiden negatiivisten ja väärin positiivisten tulosten summaan. Analyttinen valikoivuus mittaa menetelmän kykyä erottaa tutkittava kohdekanta muiden, kohdekantaa muistuttavien kantojen joukosta.

- tarkkuus
Kahden eri menetelmän samoista näytteistä antamien yhdenmukaisten tuloksien osuus suhteessa kaikkiin saatuihin tuloksiin.
- menetelmien välinen yhdenmukaisuus
Tarkkuuden avulla laskettu kahden eri menetelmän välinen yhdenmukaisuuden aste.
- toistettavuus
Kuvaa lyhytaikaista toistotarkkuutta, jossa sama henkilö saa saman tuloksen eri analyysikerroilla samoissa olosuhteissa.
- uusittavuus
Kuvaa pitkäaikaista toistotarkkuutta, jossa eri henkilöt saavat saman tuloksen samoissa olosuhteissa.

Kasvivirusten tunnistamista vaikeuttavat piilevät tartunnat, kahden tai useamman viruslajin aiheuttama sekainfektio ja virusten esiintyminen uusilla maantieteellisillä alueilla tai epätasaisesti ja pieninä määrinä kasveissa (López ym. 2003, Martin ym. 2000). Kasvivirukset saattavat muuntua esimerkiksi mutaation tai rekombinaation seurauksena ja näin kehittää uusia ominaisuuksia (Roossinck 1997). Tämä voi vaikeuttaa tunnistamista mikäli se on perustunut ominaisuuteen, joka on mutaation tai rekombinaation myötä muuttunut. Kasvivirusten luotettava ja varhainen tunnistaminen edellyttää herkkien tunnistusmenetelmien lisäksi huolellista suunnittelua miten ja mistä kasveista näytteenotto suoritetaan sekä miten näytettä käsitellään, jotta tulokset kuvaavat tilastollisesti riittävällä varmuudella taudin esiintymistä tutkittavassa populaatiossa (López ym. 2003).

2.2.1 Indikaattorikasvien käyttö kasvivirusten tunnistamisessa

Kasvivirusten tunnistaminen indikaattorikasveilla perustuu virukselle tunnusomaisten oireiden havaitsemiseen. Menetelmässä kasvit tartutetaan viruksella hieromalla virusta sisältävää kasvimehua ja hiontajauhetta lehtipinnoille, varttamalla tai altistamalla kasvit virusta kantaville vektoreille (Hull 2014, s. 756). Indikaattorikasvien käyttö sisältää useita eri rajoitteita, sillä menetelmä on työläs sekä tilaa ja aikaa vievä (Ward ym. 2004). Saman viruksen eri kannat eivät välttämättä aiheuta indikaattorikasveihin näkyviä oireita, jolloin ne jäävät tunnistamatta. Muut patogeenit ja ympäristön stressitekijät voivat aiheuttaa indikaattorikasveille samankaltaisia oireita kuin tutkittava virus, mikä vaikeuttaa tunnistamista. Menetelmän rajoitteista huolimatta viruksen biologiseen aktiivisuuteen perustuvat tunnistusmenetelmät, kuten indikaattorikasvit, ovat tärkeä osa kasvivirusdiagnoosiikkaa, sillä niiden avulla voidaan arvioida viruksen infektiokykyä (Hull 2014, s. 756).

2.2.2 Serologiset tunnistusmenetelmät

Serologiset tunnistusmenetelmät perustuvat antigeenin ja vasta-aineen välisen vuorovaikutuksen havaitsemiseen. Vasta-aineet ovat selkärankaisten eläinten immuunijärjestelmän tuottamia proteiineja, jotka kykenevät sitoutumaan elimistölle vieraisiin proteiineihin, antigeeneihin. Kasvivirusten proteiinikuori toimii tyypillisesti antigeeninä, mikä aiheuttaa koe-eläimessä immuunipuolustuksen vasteen mahdollistaen siten kasvivirusta tunnistavan vasta-aineen tuottamisen ja eristämisen. (Langham ja Brown 2016, s. 68). Serologisten tunnistusmenetelmien käyttö kasvivirusdiagnostikassa yleistyi 1970-luvulla ELISA-testin (entsyymivälitteinen immunosorbenttimääritys) kehittämisen myötä (Clark ja Adams 1977). ELISA-menetelmä on kustannustehokas ja helposti muunneltavissa, mistä johtuen se on PCR-menetelmän ohella nykyisin yksi yleisimmin käytetyistä tunnistusmenetelmistä kasvivirusdiagnostiikassa (Boonham ym. 2014).

ELISA-testin rajoitteena on, ettei vasta-aineita sisältävä antiseerumi kykene aina tunnistamaan tai erottamaan läheisesti toisilleen sukua olevia viruskantoja. Tämä voi johtua proteiinikuorta koodaavien geenien konservoituneista alueista, minä seurauksena vasta-aineet eivät kykene erottamaan eri viruskantoja toisistaan. Toinen mahdollinen syy voi olla myös eri viruskantojen välillä esiintyvät vaihtelevuudet proteiinikuoren rakenteessa, jolloin osa tutkimuksen kohteena olevista viruskannoista jää ELISA-testillä tunnistamatta (Boonham ym. 2014).

2.2.3 Polymeraasiketjureaktio (PCR)

Polymeraasiketjureaktio (PCR) on entsyymaattinen reaktio, jonka avulla haluttua DNA-jaksoa voidaan monistaa eksponentiaalisesti (Mullis ym. 1986). PCR-menetelmää on käytetty kasvitautien diagnostiikassa 1990-luvulta lähtien (Henson ja French 1993). Menetelmä perustuu DNA-polymeraasientsyymin kykyyn syntetisoida reaktioseoksessa olevista deoksiribonukleotiditrifosfaateista (dNTP), adeniinista (A), guaniinista (G), tymiinistä (T) ja sytosiinista (C), uusi komplementaarinen DNA-juoste templaattijuosteen rinnalle (Mullis ym. 1986).

DNA-polymeraasi ei kykene syntetisoimaan uutta juostetta ilman alukkeita, jotka toimivat DNA:n kahdentumisen aloituskohtana (Mullis ym. 1986). Alukkeet ovat lyhyitä, tyypillisesti synteettisesti valmistettuja, noin 20–30 nukleotidin mittaisia yksijuosteisia oligonukleotideja (Mullis 1990). PCR-reaktion aikana alukkeet sitoutuvat komplementaarisesti templaatti-DNA:n vastinjuosteiden vastakkaisiin päihin adeniniin (A) sitoutuessa tymiiniin (T) ja guaniiniin (G) sytosiiniin (C).

DNA-polymeraasi liittää alukkeen 3'-päähen reaktioseoksessa olevia nukleotideja templaatin mallin mukaan, kunnes molemmille templaatin juosteille on syntynyt vastinjuoste. Lopputuloksena alukkeiden väliin jäävä DNA-jakso monistuu. PCR:ssä saadaan monistettua tietty DNA-jakso, kun alukkeiden emäsjärjestys suunnitellaan komplementaariseksi monistettavan DNA-jakson päiden sekvenssien kanssa (Mullis ym. 1986).

PCR-reaktiossa vuorottelevat denaturointivaihe, alukkeiden kiinnittymisvaihe ja pidennysvaihe, jotka kaikki vaativat toimiakseen tietyn lämpötilan. Jokaisen syklin jälkeen monistettavan DNA-jakson määrä on kaksinkertaistunut. Denaturointivaiheen tarkoituksena on saada templaatti-DNA:n vastinjuosteet irtoamaan toisistaan. Kiinnittymisvaiheessa lämpötila säädetään optimaaliseksi alukkeiden sitoutumiselle templaatti-DNA:han ja pidennysvaiheessa lämpötila säädetään optimaaliseksi DNA-polymeraasin toiminnalle (Mullis ym. 1986).

DNA-polymeraasi ei kykene kopioimaan RNA:ta. Mikäli tutkittavan näytteen genomi on RNA:ta, siitä syntetisoidaan ensin vastin-DNA (complementary DNA, cDNA) käänteiskopioijaentsyymien avulla (Dietzgen 2002).

DNA-polymeraasin, DNA-templaatin, alukkeiden ja deoksiribonukleotiditriphosfaattien lisäksi PCR-reaktio tarvitsee toimiakseen kahdenarvoisia kationeja ja puskuriliuosta. Puskuriliuoksen tehtävänä on tarjota optimaaliset olosuhteet DNA-polymeraasin toiminnalle ja alukkeiden sitoutumiselle templaatti-DNA:han. Kahdenarvoiset kationit, kuten magnesium (Mg^{2+}), toimivat kofaktoreina DNA-polymeraasille (van Pelt-Verkuil ym. 2008). Näiden lisäksi PCR:ssä voidaan mahdollisesti käyttää erilaisia apuaineita, kuten glyserolia, betaiinia tai dimetyylisulfoxidia (DMSO), jotka helpottavat DNA-templaatin denaturaatiota (Markoulatos ym. 2002).

PCR-menetelmästä on olemassa monia eri sovelluksia. Sisäkkäisessä PCR:ssä käytetään kahta alukeparia yhden alukeparin sijaan, mikä kasvattaa menetelmän valikoivuutta. Ensimmäistä alukeparia käytetään monistamaan tiettyä DNA-jaksoa, minkä jälkeen toista alukeparia käytetään monistamaan tämän jakson sisällä olevaa sekvenssialuetta (Dietzgen 2002). Reaaliaikaisessa PCR:ssä (qPCR) monistuvan tuotteen määrää seurataan reaaliajassa fluoresoivan merkkiaineen avulla (Mackay ym. 2002).

PCR-reaktiossa monistunut tuote voidaan todeta ajamalla DNA-näyte geielektroforeesissa. Elektroforeesin toiminta perustuu agarosi- tai akryyliamidigeelissä

kulkevaan sähkövirtaan, joka erottelee nukleiinihappomolekyylit koon mukaan. Ajon aikana pienikokoiset molekyylit kulkeutuvat geelissä pidemmälle kuin suurikokoiset molekyylit. Elektroforeesissa eroteltujen nukleiinihappomolekyylien visualisointiin voidaan käyttää erilaisia väriaineita, jotka sitoutuvat nukleiinihappopoketjuihin. Nukleiinihappoihin sitoutuneiden värimolekyylien havaitsemiseen käytetään usein ultraviolettivaloa. Geeliajon perusteella voidaan päätellä vastaako PCR:ssä monistunut tuote tutkittavan näytteen DNA-molekyylin kokoa (Dietzgen 2002, Narayanasamy 2010).

PCR on herkkä, nopea ja valittavien alukkeiden avulla helposti muunneltavissa, mistä johtuen se on yleisesti käytetty menetelmä eri biologian aloilla. PCR:ssä halutun DNA-jakson monistamiseen tarvitaan templaatti-DNA:ta vain piko- tai femtogrammoja. PCR on yleensä 10^2 – 10^5 kertaa herkempi menetelmä kuin ELISA-testi. PCR-menetelmän rajoitteita ovat sen herkkyys kontaminaatioille, tarve tietää monistettavan templaatti-DNA:n sekvenssi ainakin osittain ja tarve suunnitella alukkeet, jotta haluttu DNA-jakso monistuu reaktion aikana (Dietzgen 2002). PCR:n herkkyyttä vähentävät DNA-polymeraasin toimintaa inhiboivat aineet. Näitä ovat esimerkiksi eri fenoliset yhdisteet tai polysakkaridit, jotka poistetaan näytteestä DNA:n puhdistamisen aikana (Henson ja French 1993, López ym. 2003).

2.2.3.1 PCR-menetelmän suunnittelu ja optimointi

PCR:n suunnittelun lähtökohtana on valita reaktiossa monistettava DNA-jakso ja tätä aluetta monistavat alukkeet. Alukkeet voidaan suunnitella tunnistamaan samaan heimoon tai sukuun kuuluvia eri viruslajeja tai tiettyä viruslajia, -kantaa tai -isolaattia (Singh 1998). Monistettava DNA-alue valitaan virusgenomin konservoituneelta alueelta, kun menetelmällä halutaan tunnistaa useita toisilleen sukua olevia viruksia. Konservoituneilla alueilla tarkoitetaan sekvenssejä, jotka ovat säilyneet samankaltaisina evoluution kuluessa ja joissa esiintyy vähän muunte-lua saman virusheimon tai -suvun sisällä. Konservoituneita sekvenssejä esiintyy monilla viruksilla alueilla, jotka koodaavat kuoriproteiinia ja RNA-polymeraasia (Dietzgen 2002, Henson ja French 1993, Rinehart ym. 2007).

Alukkeiden suunnittelussa on otettava huomioon etteivät ne saa sisältää itsel-leen, muille alukkeille tai muihin kuin suunniteltuun kohtaan templaatti-DNA:ssa komplementaarisia jaksoja. Alukkeiden epäspesifinen sitoutuminen lisää epäspesifisten monistustuotteiden määrää ja heikentää halutun DNA-jakson monis-tumistehokkutta (van Pelt-Verkuil ym. 2008). Optimiin lämpötilaan alukkeiden kiinnittymisvaiheelle vaikuttaa alukkeiden sulamislämpötila (T_m), jolla tarkoi-

tetaan lämpötilaa, jossa puolet DNA:sta on avautunut yksijuosteiseksi. Liian alhainen lämpötila alukkeiden kiinnittymisvaiheelle lisää epäspesifisten monistustuotteiden syntyä ja liian korkea lämpötila vähentää alukkeiden sitoutumista templaatti-DNA:han. Tämän johdosta alukkeiden poikkeavat sulamislämpötilat voivat heikentää PCR:n tehokkuutta. Alukkeiden sulamislämpötilaan vaikuttavat alukkeiden pituus ja guaniinin ja sytosiinin osuus (GC-%) emäksistä. Guaniinin ja sytosiinin välille muodostuva kovalenttinen sidos on kestävämpi kuin adeniinin ja tyrosiinin välille muodostuva sidos, minkä vuoksi alukkeen korkea GC-pitoisuus nostaa alukkeen sulamislämpötilaa. Alukkeiden suunnittelun avuksi on olemassa useita eri tietokoneohjelmia (van Pelt-Verkuil ym. 2008).

PCR-menetelmän optimoinnin tavoitteena on saada haluttua DNA-jaksoa monistumaan mahdollisimman paljon reaktion aikana ja välttää epäspesifisten monistustuotteiden syntyä (Hyndman ja Mitsuhashi 2003). Monistettavan DNA-templaatin valinnan ja alukkeiden suunnittelun jälkeen PCR-menetelmää voidaan optimoida säätämällä eri reagenssien pitoisuuksia, reaktiossa käytettäviä lämpötiloja sekä eri syklien määrää ja aikaa (Grunenwald 2003). Useat tekijät vaikuttavat PCR-menetelmän toimivuuteen, herkkyYTEEN ja valikoituvuuteen, minkä vuoksi uuden PCR-protokollan optimointi on tehtävä empiirisesti (Henson ja French 1993).

2.2.3.2 Degeneroitujen alukkeiden käyttö kasvivirusdiagnostiikassa

Kasvivirukset ovat genomiltaan hyvin vaihtelevia, mistä johtuen samaan sukuun tai heimoon kuuluvien kasvivirusten laaja-alaisessa tunnistamisessa käytetään degeneroituja alukkeita (James ym. 2006). Degeneroiduilla alukkeilla tarkoitetaan sekoitusta erilaisia sekvenssejä, jotka sisältävät eri nukleotideja samassa kohtaa sekvenssiä. Esimerkki degeneroidusta alukkeesta on GG[C,G]A[C,G,T]A, jossa kolmas emäs voi olla sytosiini tai guaniini ja viides emäs sytosiini, guaniini tai tymiini (Linhart ja Shamir 2005). Degeneroitujen alukkeiden käytöllä pyritään siihen, että ainakin osa alukesekvensseistä olisi komplementaarinen templaatti-DNA:lle. Degeneroitujen alukkeiden avulla on mahdollista tunnistaa samaan heimoon tai sukuun kuuluvia entuudestaan tuntemattomia viruksia (James ym. 2006).

Zheng ym. (2010) muistuttavat, etteivät degeneroidut alukkeet kykene välttämättä tunnistamaan kaikkia samaan heimoon tai sukuun kuuluvia viruksia. Degeneroitujen alukkeiden käytön rajoitteena on, että templaatti-DNA:lle komplementaaristen alukesekvenssien pitoisuus voi jäädä pieneksi, mikä heikentää PCR-reaktion tehokkuutta. Suppean sekvenssiaineiston perusteella suunnitellut degeneroidut alukkeet

eivät välttämättä kata kaikkia sekvenssivariaatioita, joita saman virussuvun- tai heimon tuntemattomissa lajeissa voi esiintyä.

Degeneroitujen alukkeiden käytöstä kasvivirusdiagnostiikassa on olemassa useita eri esimerkkejä (Dietzgen 2002). Esimerkiksi Zheng ym. (2010) suunnittelivat degeneroidut alukkeet (NIb2F/NIb3R) tunnistamaan suurimman kasvivirusluvun muodostavia potyvirusia. He osoittivat tekemässään tutkimuksessa NIb-alukkeiden kyetneen tunnistamaan potyvirusia kaikista tutkimuksessa käytetyistä 23:sta eri potyviruslajista ja kolmesta mahdollisesti uudesta potyviruslajista.

Untiveros ym. (2010) havaitsivat tekemässään tutkimuksessa alukkeiden (Ilar1F5/Ilar1R7) tunnistaneen laajasti useita *Bromoviridae*-heimoon kuuluvia viruslajeja. Alukkeet tunnistivat kaikki testissä käytetyt kuusi *Ilarvirus*-sukuun kuuluvaa viruslajia, sinimailasen mosaiikkiviruksen (AMV) *Alfamovirus*-suvusta, pelargonin vyöhykelaikkuviruksen (PZSV) *Anulavirus*-suvusta, kattaran mosaiikkiviruksen (BMV) *Bromovirus*-suvusta ja kolme *Cucumovirus*-sukuun kuuluvaa viruslajia. Alukkeiden todettiin ajoittain monistaneen epäspesifisiä tuotteita terveistä kasveista, mutta näiden koot eivät yleensä vastanneet *Bromoviridae*-heimolle spesifisen kokoista tuotetta (300 emäsparia).

Secoviridae-heimoon kuuluvat nepovirukset jaetaan niiden genomien rakenteen, sekvenssien samankaltaisuuden ja serologian perusteella kolmeen eri alaryhmään, a, b ja c (Kansainvälinen virusluokituskomitea 2017). Nepovirusten geneettinen vaihtelevuus vaikeuttaa nepovirusluvulle spesifisten degeneroitujen alukkeiden suunnittelua, mistä johtuen Digiaron ym. (2007) mukaan degeneroituja alukkeita on suunniteltu tunnistamaan suvun sijaan jotain tiettyä nepoviruksen alaryhmää. Wei ja Clover (2008) suunnittelivat kaksi alukeparia, joista ensimmäinen alukepari suunniteltiin tunnistamaan nepoviruksen alaryhmää a ja toinen alukepari nepoviruksen alaryhmää b. Tutkimuksessa alukkeet tunnistivat kaikki 17 nepovirusisolaattia viidestä alaryhmään a kuuluvasta lajista ja kolmesta alaryhmään b kuuluvasta lajista.

2.2.3.3 Multiplex-PCR:n käyttö kasvivirusdiagnostiikassa

Multiplex-PCR:ssä käytetään samassa reaktiossa kahta tai useampia eri DNA-jaksoille spesifisiä alukkeita. Menetelmällä voidaan monistaa useita eri DNA-jaksoja ja näin tunnistaa samalla testillä eri heimoihin ja sukuihin kuuluvia viruksia (James 1999, Menzel ym. 2002, Nassuth ym. 2000). Multiplex-PCR säästää tavalliseen PCR:n verrattuna aikaa, reagensseja ja kustannuksia, koska tällöin

jokaiselle alukeparille ei tarvitse tehdä omaa reaktioseosta (James ym. 2006). Tunnistusmenetelmän kyvystä tunnistaa useita eri viruksia samalla testillä on erityisesti hyötyä tapauksissa, joissa kasveissa esiintyy kahden tai useamman viruslajin aiheuttama sekainfektiio (López ym. 2003). Esimerkkinä multiplex-PCR:n hyödyistä kasvivirusdiagnoosissa voidaan pitää Royn ym. (2010) tekemää tutkimusta, jossa kehitettiin multiplex-PCR-testi tunnistamaan sitruksen suruviruksen (CTV) eri genotyyppisiä. Multiplex-PCR:n kuudesta alukeparista yksi oli suunniteltu tunnistamaan yleisesti kaikkia viruksen isolaatteja ja loput viisi tiettyä viruksen genotyyppiä. Tutkimuksessaan he havaitsivat 125:n virusisolatin 175:stä virusisolaattinäytteestä sisältävän kahden tai useamman sitruksen suruviruksen eri genotyypin sekainfektiota.

Haasteena multiplex-PCR:ssä on optimoida reaktio-olosuhteet kaikille alukepareille sopiviksi (López ym. 2003). Usean eri alukeparin käyttö samassa reaktiossa lisää todennäköisyyttä epäspesifisille monistumiselle. Tähän yleisenä syynä on eri alukkeiden sitoutuminen 3'-päistä toisiinsa, jolloin ne muodostavat alukedimeerejä (Elnifro ym. 2000). Yleisenä ilmiönä multiplex-PCR:ssä on kohdesekvenssien epätasainen monistuminen, jolloin tehokkaammin monistuvat kohdesekvenssit käyttävät enemmän reaktioseoksen reagensseja monistumiseensa jättäen vähemmän reagensseja muiden kohdesekvenssien monistumiseen (Markoulatos ym. 2002). Kohdesekvenssien epätasaista monistumista esiintyi esimerkiksi Tuon ym. (2014) tekemässä tutkimuksessa, missä ongelma ratkaistiin alentamalla tehokkaasti monistavien ja kasvattamalla heikosti monistavien alukeparien pitoisuutta. Markoulatoksen ym. (2002) mukaan usean eri alukeparin käyttö samassa reaktiossa lisää alukkeiden välistä kilpailua reaktioseoksen reagensseista, mistä johtuen esimerkiksi magnesiumin, dNTP:n tai polymeraasientsyymin korkeamman konsentraation käyttö reaktioseoksessa kuin mitä tyypillisesti käytetään simplex-PCR:ssä voi lisätä multiplex-PCR:n herkkyyttä ja tehokkuutta. Esimerkiksi Gambinon ja Gribaudon (2006) tekemässä tutkimuksessa multiplex-PCR saatiin toimivammaksi nostamalla PCR-puskurin, dNTP:n ja *Taq* DNA-polymeraasientsyymin konsentraatiota korkeammaksi kuin mitä kirjallisuudesta etsittyjen alukkeiden alkuperäisissä simplex-PCR ohjeissa oli käytetty. Elnifron ym. (2000) mukaan ennen kuin multiplex-PCR:ää voidaan käyttää kliinisessä tutkimuksessa, on jokaisen alukeparin toimivuus, herkkyys ja valikoivuus testattava erikseen ja yhdessä sekä kohde-DNA:sta tehdyillä laimennoksilla että kliinisillä näytteillä.

Monissa multiplex-PCR-testeissä on käytetty kahdesta viiteen eri alukeparia (Grieco ja Gallitelli 1999, Pallás ym. 2018, Saade ym. 2000, Sharman ym. 2000), mutta vain harva tutkimus on osoittanut yli viiden eri alukeparin toimineen samassa

multiplex-PCR-testissä onnistuneesti (Gambino ja Gribaudo 2006, Pallás ym. 2018, Sanchez-Navarro ym. 2005). Multiplex-PCR:ssä monistuvien DNA-jaksojen koot on oltava riittävän erilaisia toisistaan, jotta tulosten varmistaminen geelielektroforeesin avulla on mahdollista (James ym. 2006). Henegariu ym. (1997) ehdottavat seuraavia vaihtoehtoja multiplex-PCR:n optimointiin:

- A. Kaikki PCR-tuotteet näkyvät heikosti
 - Kasvata alukkeiden pidennysvaiheen kestoa.
 - Alenna alukkeiden pidennys- ja /tai kiinnittymisvaiheen lämpötilaa.
 - Kokeile yhdistelmiä yllä olevista vaihtoehdoista.
- B. Lyhyet PCR-tuotteet (100–300 nt) näkyvät heikosti
 - Kasvata puskurin KCl-pitoisuutta.
 - Alenna alukkeiden pidennys- ja /tai kiinnittymisvaiheen lämpötilaa.
 - Kasvata alukkeiden pitoisuutta reaktioseoksessa.
 - Kokeile yhdistelmiä yllä olevista vaihtoehdoista.
- C. Epäspesifisiä PCR-tuotteita on näkyvissä
 - Kasvata puskurin KCl-pitoisuutta pitkille PCR-tuotteille.
 - Alenna puskurin KCl-pitoisuutta lyhyille PCR-tuotteille.
 - Nosta alukkeiden kiinnittymisvaiheen lämpötilaa.
 - Nosta reaktioseoksen Mg^{2+} -pitoisuutta.
 - Kokeile yhdistelmiä yllä olevista vaihtoehdoista.
- D. Pitkät PCR-tuotteet (yli 300 nt) näkyvät heikosti
 - Kasvata alukkeiden pidennysvaiheen kestoa.
 - Nosta alukkeiden pidennys- tai kiinnittymisvaiheen lämpötilaa.
 - Kasvata alukkeiden pitoisuutta reaktioseoksessa.
 - Vähennä puskurin KCl-pitoisuutta.
 - Kokeile yhdistelmiä yllä olevista vaihtoehdoista.
- E. Jos A-, B-, C- tai D-vaihtoehdot eivät toimi
 - Käytä toista DNA-preparaattia.
 - Vaihda reagenssit tuoreisiin reagensseihin.
 - Suunnittele alukkeet uudestaan.

2.2.4 Syväsekvensoinnin käyttö kasvivirusdiagnostiikassa

Sekvensoinnin tarkoituksena on tuottaa tutkittavasta DNA:sta sitä vastaava tietokoneella käsiteltävä merkkijono. Syväsekvensointi (HTS, high throughput sequencing) tarkoittaa menetelmää, jonka avulla miljoonia DNA-molekyylejä voidaan sekvensoida samanaikaisesti (Massart ym. 2014).

Syväsekvensoinnin suorittaminen koostuu useista vaiheista, jotka voidaan jakaa templaatin valmistukseen, sekvensointiin, kuvantamiseen ja data-analyysiin. Bioinformatiikan keinoin voidaan vertailla sekvenssien emäsjärjestyksiä ja niissä olevien päällekkäisyyksien avulla rakentaa viruksen perimä osittain tai kokonaan (Metzker 2010).

Adamsin ja Foxin (2016) mukaan syväsekvensointia käytettiin kasvivirusdiagnoosiin ensimmäisen kerran vuonna 2009 kolmen eri tutkimusryhmän toimesta (Adams ym. 2009, Al Rwahnih ym. 2009, Kreuze ym. 2009). Syväsekvensoinnin avulla on pystytty tunnistamaan useita, myös entuudestaan tuntemattomia, viruksia (Coetzee ym. 2010, Jo ym. 2016, Kashif ym. 2012). Syväsekvensoinnin suosio on kasvanut kasvivirusdiagnoosiin ja se on kehittymässä yhdeksi rutiinidiagnostiikan välineistä (Pecman ym. 2017).

Kasvivirusten sekvensointiin on kehitetty erilaisia keinoja, joiden merkittävimmät eroavaisuudet ovat millaisiin nukleiinihapporakenteisiin sekvensointi kohdistuu ja miten sekvensointi on pyritty kohdentamaan viruksista peräisin oleviin nukleiinihappoihin isäntäkasvista olevien nukleiinihappojen sijaan (Boonham ym. 2014). Eräs kasvivirusdiagnoosikeino on sekvensoida kaikki näytteessä esiintyvä lähetti-RNA (mRNA), millä on onnistuneesti tunnistettu useita eri viruksia (Al Rwahnih ym. 2009, Wylie ja Jones 2011, Wylie ym. 2012). Menetelmän rajoitteena on, että suurin osa sekvensoitavasta lähetti-RNA:sta on peräisin isäntäkasvista, minkä seurauksena alhaisena pitoisuutena esiintyvät virukset voivat jäädä tunnistamatta (Boonham ym. 2014). Toinen sekvensointikeino kasvivirusdiagnoosiin on sekvensoida puhdistettua kaksijuosteista RNA:ta (dsRNA:ta) (Coetzee ym. 2010). Jotkin viruksista, kuten negatiiviset yksijuosteiset RNA-virukset (-ssRNA), tuottavat vain vähän kaksijuosteista RNA:ta, mistä johtuen menetelmä ei sovellu kaikkien virusten tunnistamiseen (Boonham ym. 2014). Sekvensointi on pyritty kohdentamaan viruksista peräisin oleviin nukleiinihappoihin esimerkiksi poistamalla näytteessä kasvista peräisin oleva ribosomaalinen RNA (rRNA) (Pecman ym. 2017) ja puhdistamalla viruspartikkelit kasvimateriaalista (Wu ym. 2015).

Syväsekvensoinnissa tapahtuneen kehityksen ansiosta on Olmoksen ym. (2018) mukaan yhä todennäköisempää, että uudet virukset ja uudet isäntäkasvit jo tunnetuille viruksille saadaan tunnistettua. Samalla he kuitenkin muistuttavat uusien viruslöydösten tuomista haasteista, mitkä liittyvät niiden biologisen merkittävyyden, kuten patogeenisuuden, isäntäkirjon tai epidemiologian, selvittämiseen. Myös Roossinck (2017) toteaa, ettei virussekvenssin löytyminen näytteestä yksinään riitä osoittamaan löydetyn viruksen olevan syynä kasvissa havaittaviin oireisiin.

Esimerkiksi jotkin kasviviruksista voivat olla integroituneena osaksi isännän genomi, jolloin virus saattaa olla inaktiivinen eikä monistu tai aiheuta oireita kasvilla (Kreuze 2014). Massart ym. (2017) tuovat esille tilanteen ongelmallisuutta toteamalla uusien virusten perusteellisen tuhoojariskianalyysin kasvinterveysstandardien ISPM 2 ja ISPM 11 (FAO 2016, FAO 2017a) mukaan vievän mahdollisesti useita vuosia. Samassa artikkelissa Massart ym. (2017) esittelevät ohjeistuksen mitä tehdä uuden viruksen löytyessä, biologisesta määityksestä tiedonantoon viranomaisille.

2.2.5 siRNA-syväsekvensoinnin käyttö kasvivirusdiagnoosissa

Kasvivirusdiagnoosissa on yleistynyt sekvensointitapa, jossa sekvensoidaan kasviin kertyneitä pieniä RNA-molekyylejä (small interfering RNA, siRNA) (Kreuze ym. 2009, Pecman ym. 2017). Menetelmää on sovellettu kasvivirusten tutkimisessa ensimmäisen kerran Kreuzen ym. (2009) toimesta. siRNA-syväsekvensoinnin käyttö kasvivirusten tunnistusmenetelmänä perustuu kasvien kykyyn suojautua virustartunnoilta RNA-hiljennyksen avulla. RNA-hiljennyksessä kasvin puolustusmekanismi tunnistaa vierasperäisen kaksijuosteisen RNA:n ja hajottaa sen 21–24 nukleotidin mittaisiksi siRNA-molekyyleiksi. siRNA-syväsekvensoinnissa 21–24 nukleotidin pituiset RNA-molekyylit eristetään kasvista ja sekvensoidaan. siRNA-molekyylien sekvenssien välillä esiintyy päällekkäisyyttä, mistä johtuen sekvensseistä voidaan koota eripituisia yhtenäisiä sekvenssejä. Vertailemalla koottuja sekvenssejä tietokannoista löytyvien virusten perimään voidaan selvittää näytteessä esiintyvät virukset ja viroidit (Kreuze 2014). Kreuze ym. (2014) esittelevät artikkelissaan siRNA-sekvenssidatan analysointiin kuuluvat tyypilliset työvaiheet, joita ovat:

1. Sekvenssien kokoaminen *de Novo* ilman referenssisekvenssiä.
2. Koottujen sekvenssien vertaaminen tietokannassa oleviin virussekvensseihin.
3. siRNA-sekvenssien linjaus eniten vastaaviin referenssisekvensseihin.
4. Alukkeiden suunnittelu ja tulosten varmistaminen PCR:llä.

Poogginin (2018) artikkeli esittelee täydellisen listan kaikista kasveista, joista on löydetty viruksia siRNA-syväsekvensoinnilla. Lista sisältää lähes 150 eri kasvilajia ja yli 300 eri viruslajia. Osassa siRNA-menetelmää koskevissa tutkimuksissa on pystytty tunnistamaan viruksen lähes koko genomi (Kutnjak ym. 2014, Li ym. 2012, Zhang ym. 2014). siRNA-menetelmällä on saatu tietoa virusten esiintymisestä ja niiden isäntäkasveista. Esimerkiksi Zhang ym. (2011) löysivät siRNA-menetelmällä DNA-viruksiin luokiteltavia badnaviruksia, mikä oli ensimmäinen havainto DNA-virusten esiintymisestä viiniköynnöksissä (*Vitis vinifera* L.). Candresse ym. (2014)

löysivät siRNA-menetelmän avulla sokeriruo'osta (*Saccharum officinarum* L.) entuudestaan tuntemattoman viruksen, jota ei ollut löydetty karanteenikasvintuhoojien varalta perinteisillä menetelmillä tehdyissä kasvintarkastuksissa.

siRNA-menetelmän merkittävimmät edut ovat, ettei se edellytä ennakko-oletusta näytteessä mahdollisesti esiintyvistä viruksista ja viroideista ja että menetelmällä on mahdollista havaita kaikki virukset ja viroidit yhdellä kertaa (Kreuze 2014). Tästä johtuen menetelmän on ehdotettu soveltuvan kasvintarkastukseen, varmennetun taimituotannon testaamiseen ja yleisesti laaja-alaiseen virusten tunnistamiseen suuresta näytemäärästä (Hadidi ym. 2016, Olmos ym. 2018). Serologiset ja perinteiset molekulaariset tunnistusmenetelmät ovat viruslaji- tai virusryhmäkohtaisia, mistä johtuen niiden suorittaminen vaatii ennakko-oletuksen näytteessä olevista mahdollisista viruksista. Tämä vaikeuttaa virusten tunnistamista tilanteissa, joissa virukset ovat entuudestaan tuntemattomia tai virus infektoi uutta isäntäkasvia (Adams ym. 2009).

Kreuzen (2014) mukaan siRNA-menetelmän herkkyys on verrattavissa PCR:ään, mistä johtuen menetelmän katsotaan soveltuvan myös pieninä pitoisuuksina esiintyvien virusten tunnistamiseen. siRNA-menetelmän herkkyys riippuu sekvensoitavien siRNA-molekyylien määrästä ja sekvenssien kokoamisessa käytettävistä algoritmeista ja parametreista (Adams ja Fox 2016, Zerbino ja Birney 2008). Pallásin ym. (2018) mukaan siRNA-menetelmä tarvitsee perusteellista validointia ja herkkyyden vertaamista esimerkiksi reaaliaikaiseen RT-PCR-menetelmään ennen kuin voidaan varmistua siRNA-menetelmän soveltuvuudesta osaksi kasvivirusien rutiinidiagnostikkaa. siRNA-menetelmän herkkyyttä puoltaa Santalan ja Valkosen (2018) tekemä tutkimus, jossa siRNA-menetelmä oli 10 kertaa herkempi verrattuna reaaliaikaiseen RT-PCR-menetelmään, kun tiedettiin etukäteen mitä viruksia näyte todennäköisesti sisältää.

siRNA-menetelmän haasteena on muiden HTS-menetelmien tavoin sekvensoinnista saatavan suuren tietomäärän käsittely (Roossinck 2017). Virusten tuottamat RNA-hiljennyksenestäjaproteiinit vähentävät siRNA-molekyylien määrä kasveissa, minkä vuoksi niiden on arvioitu heikentävän siRNA-menetelmän kykyä tunnistaa viruksia näytteestä (Roossinck ym. 2015). Toisaalta menetelmällä on löytynyt kaikkiin kasvivirusryhmiin kuuluvia viruksia useista kasviheimoista (Pooggin 2018).

Tuntemattoman viruksen tunnistaminen siRNA-menetelmällä onnistuu, mikäli näytteessä esiintyvän viruksen sekvenssijärjestys on riittävän samankaltainen

jonkin tietokannassa olevan virussekvenssin kanssa. Tästä johtuen virukset, joiden nukleotidi- tai aminohapposekvenssijärjestys vastaa vähän tai ei lainkaan tunnettuja virussekvenssejä, voivat jäädä menetelmällä tunnistamatta. Tunnetut virukset ovat genomiltaan hyvin vaihtelevia, mistä johtuen on syytä epäillä olevan tuntemattomia viruksia, jotka eivät vastaa genomiltaan tunnettuja virusheimoja tai -sukuja (Kreuze 2014). Tietoa eri virusgenomeista karttuu jatkuvasti tietokantoihin, mistä johtuen samoista näytteistä voidaan mahdollisesti löytää myöhemmin uusia viruksia (Roossinck 2012).

Kasvista peräisin olevat siRNA-sekvenssien pituutta vastaavat pikku-RNA-molekyylit voivat vähentää algoritmien kykyä koota viruksista peräisin olevista siRNA-sekvensseistä yhtenäisiä sekvenssejä. Kasvien genomien määrittäminen tehostaa siRNA-menetelmän kykyä tunnistaa viruksia, koska tällöin kasvista peräisin olevat pikku-RNA-molekyylit voidaan poistaa sekvenssiaineistoista ennen sekvenssien kokoamista. Menetelmä ei suoraan auta tunnistamaan tuntemattomia viruksia, mutta tällöin koottujen, yhtenäisten siRNA-sekvenssien voidaan epäillä olevan peräisin viruksista (Kreuze 2014).

2.2.5.1 Virustartunnan aktivoima RNA-hiljennys

RNA-hiljennys on evolutiivisesti konservoitunut ja useimmilla eukaryoottisilla eliöillä esiintyvä geenin hiljennysmekanismi, jossa kaksijuosteinen RNA toimii sekä hiljennuksen kohteena että sen käynnistäjänä. RNA-hiljennyksellä on tärkeä osa kasvisolun toimintojen säätelyn lisäksi kasvin immuunijärjestelmän kyvyssä puolustautua virustartuntoja vastaan (Baulcombe 2004, Wang ym. 2012). Virustartunnan aktivoima RNA-hiljennys perustuu siihen, että kasvivirus tuottavat jossakin vaiheessa monistumisena aikana kaksijuosteista RNA:ta tai kaksijuosteisia RNA-rakenteita, jotka tunnistetaan ja hajotetaan. Hajoamistuotteiden (siRNA) avulla puolestaan tunnistetaan virusperäistä yksijuosteista RNA:ta, joka siten myös altistuu hajotukselle. Tällä tavoin kasvi voi estää viruksen monistumista ja parantua virustartunnan aiheuttamista oireista (Mlotshwa ym. 2008).

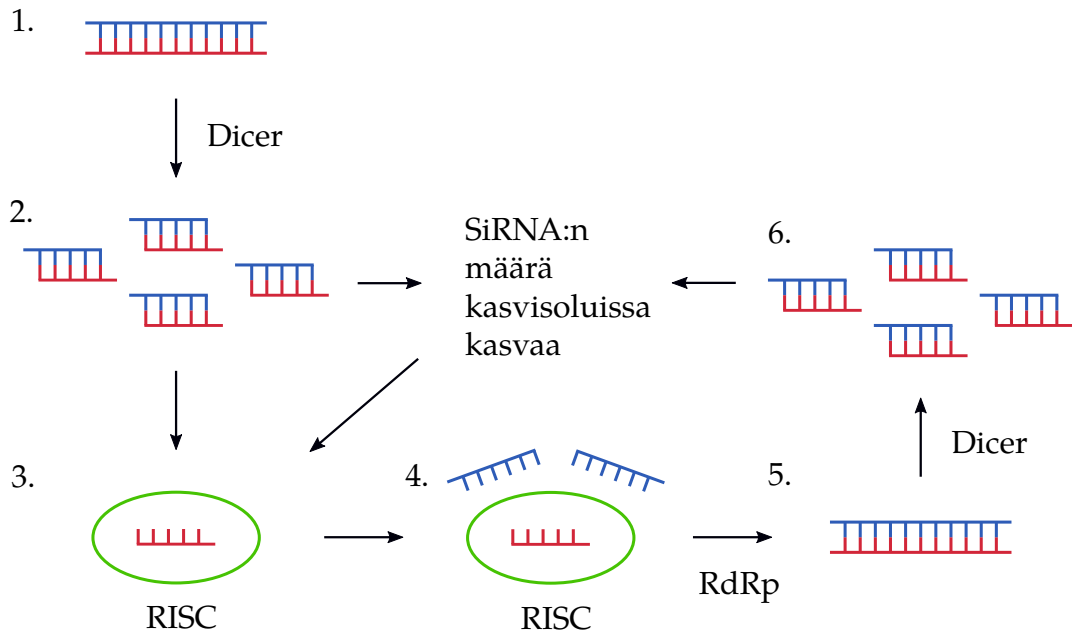
Virustartunnan aktivoima RNA-hiljennys käynnistyy (kuva 1), kun kasvisolun DICER-entsyymit (DCL) tunnistavat ja hajottavat vierasperäistä kaksijuosteista RNA:ta, tai yksijuosteista, mutta hiuspinnimäisiä rakenteita muodostavaa RNA:ta 21–24 nukleotidin mittaisiksi siRNA-molekyyleiksi. Hiuspinnimäisiä rakenteita muodostavaa RNA:ta syntyy saman RNA-molekyylin emästen pariutuessa keskenään. (Baulcombe 2004, Hamilton ja Baulcombe 1999).

DCL-entsyymit kuuluvat kasvisolun kolmostyyppin endoribonukleaaseihin, jotka ovat erikoistuneet kaksijuosteisen RNA:n hajotukseen (Soosaar ym. 2005). Lituruoholla (*Arabidopsis thaliana* L.) tiedetään esiintyvän neljä erilaista DCL-entsyymiä, jotka osallistuvat suoraan tai välillisesti siRNA-molekyylien tuotantoon (Mlotshwa ym. 2008). Blevinsin ym. (2006) mukaan DCL1-entsyymi ei vaikuta olevan suoraan yhteydessä siRNA-molekyylien tuotantoon RNA-viruksista, mutta se voi edistää 21:n nukleotidin mittaisten siRNA-molekyylien tuotantoa DNA-viruksista. DCL2 ja DCL4-entsyymit osallistuvat siRNA-molekyylien tuotantoon RNA-viruksista (Ding ja Voinnet 2007). DCL3-entsyymit ovat tärkeitä kasvin puolustuksessa DNA-viruksia vastaan ja entsyymien tuottamien siRNA-molekyylien on todettu esimerkiksi osallistuvan DNA:n metylaatioon (Chan ym. 2004).

Kaksijuosteisen RNA:n hajotuksen jälkeen yksijuosteista RNA:ta hajottava reaktioketju syntyy, kun ARGONAUTE-proteiinit (AGO) sitovat siRNA-molekyylistä yhden juosteen ja hajottavat tämän vastinjuosteen. AGO-proteiini ja tähän sitoutunut siRNA-juoste toimivat RISC-kompleksina (RNA-induced silencing complex). RISC-kompleksi tunnistaa ja katkoo sille komplementaarista yksijuosteista RNA:ta, mikä johtaa viruksen muuttumiseen toimintakyvyttömäksi (Baulcombe 2004).

Lituruoholla on todettu esiintyvän 10 ja riisillä (*Oryza sativa* L.) 19 AGO-proteiinia koodaavaa geeniä (Carbonell ja Carrington 2015, Nonomura ym. 2007). AGO-proteiinien toiminta on vielä osittain tuntematon, sillä Carbonellin ja Carringtonin (2015) mukaan esimerkiksi AGO-proteiinien toimintaa ohjaavien molekulaaristen signaalien toiminta ei ole täysin selvitetty.

Kasvit kykenevät vahvistamaan RNA-hiljennystä tuottamalla kaksijuosteista RNA:ta vierasperäisestä RNA:sta. Kasvin RNA:sta riippuvainen RNA-polymeraasi (RdRp) rakentaa yksijuosteiselle RNA:lle vastinjuosteen. Muodostunut kaksijuosteinen RNA tunnistetaan ja hajotetaan DICER-entsyymien avulla. Tapahtumassa siRNA-molekyylien määrä kasvaa nopeasti, minkä seurauksena yhä useampi AGO-proteiini voi sitoa virusgenomia vastaavan siRNA-juosteen (Mlotshwa ym. 2008).



Kuva 1. siRNA-molekyylien lähteet jaetaan primäärisiin ja sekundäärisiin lähteisiin. siRNA-molekyylien synty kasvisoluissa jaetaan seuraaviin vaiheisiin: 1) Vierasperäinen kaksijuosteinen RNA. 2) DICER-entsyymi tunnistaa ja hajottaa kaksijuosteista RNA:ta. Hajotuksessa muodostuu primäärisiä siRNA-molekyyliä 3) AGO-proteiini sitoo siRNA-molekyylistä yhden juosteen ja hajottaa tämän vastinjuosteen. AGO-proteiini ja tähän sitoutunut siRNA-juoste muodostavat RISC-kompleksin. 4) RISC-kompleksi tunnistaa ja katkoo sille komplementaarista yksijuosteista RNA:ta. 5) RdRp rakentaa yksijuosteiselle vierasperäiselle RNA:lle vastinjuosteen. 6) DICER-entsyymi tunnistaa ja hajottaa kaksijuosteista RNA:ta. Hajotuksessa syntyy sekundäärisiä siRNA-molekyyliä (Baulcombe 2004).

2.2.5.2 Virusten puolustautuminen RNA-hiljennystä vastaan

Kasvivirukset ja niiden isännät käyvät evolutiivista kilpavarustelua isäntien kehittäessä uusia vastustusstrategioita viruksia vastaan ja virusten sopeutuessa näihin uusiin puolustusmekanismeihin. Kasviviruksille on kehittynyt evoluution kuluessa useita keinoja, joiden avulla ne voivat välttää RNA-hiljennystä (Baulcombe 2004). Monien kasvivirusten genomien on todettu sisältävän yhtä tai useampia hiljennyksenestäjiä koodaavia proteiineja, jotka estävät tai häiritsevät RNA-hiljennyksen eri vaiheita (Wang ym. 2012).

Virusten RNA-hiljennyksenestäjät voivat sitoutua siRNA-molekyyliin, mikä estää RISC-kompleksin muodostumista ja näin yksijuosteista RNA:ta hajottavan reaktioketjun syntyä (Lakatos ym. 2006). RNA-hiljennyksenestäjät voivat myös häiritä RNA-hiljennykselle tärkeiden proteiinien toimintaa esimerkiksi hajottamalla AGO1-proteiinia (Baumberger ym. 2007).

2.3 Kasvivirusten esiintyminen luonnonvaraisissa kasveissa

Kasvivirusten esiintymisestä luonnonvaraisissa kasveissa tiedetään vähän, sillä tutkimukset ovat keskittyneet viljelykasveissa esiintyviin virustauteihin. Osaltaan luonnonvaraisissa kasveissa esiintyvien virusten tutkintaa on vaikeuttanut yleisten tunnistusmenelmien, kuten ELISA:n ja PCR:n, vaatimus ennako-oletuksesta näytteessä esiintyvistä viruksista (Jones 2009).

Luonnonvaraiset kasvit voivat toimia kasvivirusten infektiolähteinä viljelykasveille, mistä johtuen tietämys virusten esiintymisestä luonnonvaraisissa kasvilajeissa voi auttaa viruksen aiheuttamien satovahinkojen torjunnassa (Cooper ja Jones 2006). Esimerkiksi Tugume ym. (2008) tutkivat, esiintyykö bataattia (*Ipomoea batatas* L.) infektoivaa virustautia, bataatin höyhenläikkävirusta (*Sweet potato feathery mottle virus*, SPFMV), bataatin viljelyalueilla esiintyneissä luonnonvaraisissa kasveissa. Tutkimuksen tulokset osoittivat 22:n eri kiertokasvien (*Convolvulaceae*) heimoon kuuluvan luonnonvaraisen kasvilajin kantavan bataatin höyhenläikkävirusta. Näistä 19 oli entuudestaan tuntemattomia viruksen isäntäkasveja.

Kasvivirusten ja niiden isäntien välillä on tapahtunut koevoluutiota pitkään ennen kasvien domestikaatiota, mistä johtuen luonnonvaraiset kasvit toimivat kasvivirusten monimuotoisuuden alkuperäisenä lähteenä (Jones 2009). Viimeaikaiset metagenomiset tutkimukset ovat osoittaneet, että vain murto-osa viruslajien monimuotoisuudesta tunnetaan (Stobbe ja Roossinck 2014).

Luonnonvaraisissa kasviyhteisöissä kasvivirusten aiheuttamien epidemioiden ajatellaan yleisesti olevan viljelykasveissa esiintyviin epidemioihin verrattuna lievempiä ja vähemmän vallitsevia. Epidemian leviämistä ja sen aiheuttamia tuhoja vähentävät vektorien saalistajat ja parasiitit sekä virukselle sopivien isäntäkasvien esiintyminen erillään toisistaan omina populaatioinaan. Koevoluution kautta syntyneen valintapaineen kautta luonnonvaraisille kasveille on voinut kehittyä resistenssi tai toleranssi joitakin viruslajeja vastaan (Jones 2009).

Luonnonvaraisia kasvilajeja käytetään tutkimuskohteina ymmärtämään kasvien ja virusten välisiä ekologisia ja evolutiivisia vuorovaikutussuhteita. Kasvivirusten ja näiden isäntien välisten monimutkaisten vuorovaikutussuhteiden ymmärtäminen voi mahdollisesti auttaa kasvivirusten torjunnassa ja leviämisen hallinnassa (Cooper ja Jones 2006, Penczykowski ym. 2016). Esimerkiksi kasvivirusten on todettu joissakin tapauksissa hyödyttävän kasvia sietämään erilaisia stressiolosuhteita. Xu ym. (2008) havaitsivat tekemässään tutkimuksessa kurkun mosaiikkiviruksen

(CMV), tupakan mosaiikkivirusten (TMV), tupakan rattleviruksen (TRV) ja kattan mosaiikkiviruksen (BMV) edistävän testikasvien kykyä sietää kuivuutta. Vastaavanlaiset tutkimukset herättävät toiveita hyödyntää esimerkiksi viljelykasveille suosiollisia viruksia, jotka voisivat joko parantaa satoa tai suojella kasveja ympäristön stressitekijöiltä (Roossinck 2015).

2.4 Kasvinterveyden ja taimiaineiston valvonta

Suomessa laki kasvinterveyden suojelemisesta (702/2003) pyrkii ylläpitämään hyvää kasvinterveyden tilaa ja siten edistämään maa-, metsä- ja puutarhatalouden ja elintarviketuotannon toimintaedellytyksiä sekä elintarvikkeiden turvallisuutta ja tuotteiden laatua. Laki koskee toimenpiteitä, joilla torjutaan kasvintuhoojia ja estetään niiden leviäminen (Valtioneuvosto 2003). Taimiaineistolain (1205/1994) tarkoituksena on turvata taimiaineiston terveys ja laatu (Valtioneuvosto 1994).

”Kasvinterveyden ja taimiaineiston valvonnan tavoitteena on estää kasvinterveyslainsäädännössä mainittujen vaarallisten kasvintuhoojien leviäminen Suomeen ja todettujen esiintymien hävittäminen. Lisäksi valvonnan tavoitteena on varmistaa, että markkinoitavat, maahantuotavat ja maasta vietävät kasvituotteet täyttävät muut kasvinterveyslain sekä taimiaineistolain asettamat vaatimukset.” (Evira 2018)

Karanteenikasvintuhoojalla ja vaarallisella kasvintuhoojalla tarkoitetaan lainsäädännössä määriteltyjä vaikeasti torjuttavia kasvitauteja ja tuhoeläimiä, jotka aiheuttavat merkittäviä taloudellisia vahinkoja maatalous-, metsätalous- tai puutarhatuotannossa. Karanteenikasvintuhoojia ei saa esiintyä tuotantopaikoilla tai myytävissä, maahan tuotavissa tai maasta vietävissä kasveissa. Tarkastuksissa etsitään vaarallisia kasvintuhoojia ja tarvittaessa näytteistä tunnistetaan kasvintuhoojat virallisissa laboratorioissa. Karanteenikasvintuhoojien valvonnasta vastaa Suomessa 1.1.2019 alkaen Ruokavirasto (Valtioneuvosto 2018). Karanteenikasvintuhoojat ovat lueteltuna EU:n kasvinterveysdirektiivissä (2000/29/EY) ja maa- ja metsätalousministeriön kasvinterveysasetuksessa (17/08) (Euroopan unionin neuvosto 2000, Maa- ja metsätalousministeriö 2008).

Esimerkki Ruokaviraston (entisen Elintarviketurvallisuusvirasto Eviran) tekemästä valvonnasta on vuonna 2014 taimitarhalta löydetystä, vaaralliseksi kasvitautiksi luokitellusta, luumun rokkoviruksesta (*Plum pox virus*, PPV). Löydös oli ensimmäinen havainto taudin esiintymisestä Suomessa (Santala ja Soukainen 2015).

3 Tutkimuksen tavoitteet

Tämän tutkimuksen tavoitteena oli optimoida Eviran (nyk. Ruokavirasto) kasvintuhoojalaboratoriolle kaksi tai useampi yksivaiheinen multiplex-RT-PCR-testi yhdeksälle alukeparille poty-, potex-, tobra- ja tospovirussukujen, tobamovirus alaryhmä yhden, nepoviruksen a- ja b-alaryhmien, *Bromoviridae*-virusheimon ja *Tombusviridae*-heimoon kuuluvien carmo-, diantho- ja tombusvirussukujen tunnistamiseen. Tavoitteena oli näin luoda menetelmä, joka mahdollistaa useiden eri RNA-virussukuihin kuuluvien virusten tunnistamisen samalla testillä. Multiplex-RT-PCR-testiä on tarkoitus käyttää RNA-virusten seulontatestinä kasvinäytteille, joissa testauksen piiriin kuuluvien virusten tartunta on oireiden tai isäntäkasvin perusteella todennäköistä tai mahdollista.

Tutkimuksessa käytettiin sekä multiplex-RT-PCR-testiä että siRNA-menetelmää selvittämään Laukaan ydinkasviaineistosta saaduista viroottisista näytteistä, viroottisilta vaikuttaneista luonnonvaraisista kasveista ja Kumpulan kasvitieteellisestä puutarhasta kerätyistä kasvinäytteistä esiintyvät kasviviruset. Tunnistamalla kasveissa esiintyvät viruset haluttiin osaltaan arvioida multiplex-RT-PCR:n soveltuvuutta kasvintuhoojalaboratorion käyttöön. siRNA-menetelmää käytettiin multiplex-PCR-testien tulosten oikeellisuuden varmistamisen. Lisäksi saatiin tietoa siRNA-menetelmän käytettävyydestä virusten testaamisessa.

Vähäinen tietämys luonnonvaraisissa kasveissa esiintyvistä viruksista (Jones 2009) ja siRNA-menetelmän kyky tunnistaa viruksia laaja-alaisesti (Kreuze ym. 2009) antoivat pohjan hypoteesille, että näytteistä voidaan tunnistaa mahdollisesti viruksia, joita ei ole aiemmin tunnistettu.

4 Tutkimusaineisto ja -menetelmät

4.1 Kasvimateriaali

Tutkimuksen kasviaineisto koostui Juha Tuomolan vuosien 2013–2015 aikana Helsingistä keräämistä viroottisilta vaikuttaneista luonnonkasvinäytteistä (kuva 2) (JT-alkuiset koodinumerot), Laukaan ydinkasviaineistosta saaduista viroottisista näytteistä (koodinumerot (787–791 ja 870–872) sekä kahdesta näytteestä Kumpulan kasvitieteellisestä puutarhasta (koodinumerot K1 ja K2) (taulukko 1). Näytteitä on säilytetty keräämiseen jälkeen pakastimessa (–20 °C) tai syväjäähdytyskammerissa (–80 °C). Kiinanjaloangervo- ja kvinoa-näytteiden epäiltiin kantavan arabiksen mosaiikkivirusta perustuen ennen tutkimuksen aloittamista tehtyihin testikasvitesteihin. Kasvivirusten esiintymistä näytteissä K1, K2 ja JT17 tutkittiin ainoastaan siRNA-syväsekvensoinnilla.

Taulukko 1. Tutkimuksessa käytetty kasvimateriaali.

Suomenkielinen nimi	Latinankielinen nimi	Koodi
Poimulehti	<i>Alchemilla</i> sp. L.	JT5
Suikeroalpi	<i>Lysimachia nummularia</i> L.	JT2
Keltamo	<i>Chelidonium majus</i> L.	JT3
Valkomesikkä	<i>Melilotus albus</i> Medik.	JT15
Metsähaapa	<i>Populus tremula</i> L.	JT16
Nokkonen	<i>Urtica dioica</i> L.	JT17
Metsäapila	<i>Trifolium medium</i> L.	JT14
Lupiini	<i>Lupinus polyphyllus</i> Lindl.	JT12
Valvatti	<i>Sonchus</i> sp. L.	JT4
Kiinanjaloangervo	<i>Astilbe chinensis</i> (Maxim.) Franch. & Sav.	787, 788,
Kvinoa	<i>Chenopodium quinoa</i> Willd.	789, 790
Alppiruusu	<i>Rhododendron</i> sp. L.	791
Kuunlilja	<i>Hosta</i> sp. Tratt.	870, 871, 872
Paatsama	<i>Rhamnus schneideri</i> H. Lév. & Vaniot	K1
Herttahuisku-unikko	<i>Macleaya cordata</i> (Willd.) R. Br.	K2



(a) Keltamo



(b) Poimulehti



(c) Suikeroalpi



(d) Valkomesikkä

Kuva 2. Malliesimerkkejä viroottisilta vaikuttaneista, tutkimusaineistoon valituista luonnonvaraisista kasveista. Kuvien julkaisemiseen on saatu lupa Juha Tuomolalta.

4.2 RNA:n eristäminen multiplex-RT-PCR-testiä varten

RNA:n eristäminen tehtiin RNeasy Plant Mini Kitillä (Qiagen, Saksa) valmistajan ohjeiden mukaan (Qiagen 2012, s. 50–53). Kaikki työvaiheet ja sentrifugoinnit tehtiin huoneenlämmössä, jotta RNA sitoutuu kunnolla pylvaiden kalvoille. RNA:n hajoamisen estämiseksi uuttopuskuri lisättiin jäisiin näytteisiin ennen niiden sulamista.

Pakastimessa (-20°C) tai syväjäähdytyskammiossa (-80°C) säilytettyjen lehtinäytteiden jauhamiseen käytettiin 2010 Geno/Grinder -kuularavistelijaa (Spex Sample-Prep, Yhdysvallat). Ennen näytteiden jauhamista kutakin näytettä varten varattiin yksi 5 ml:n kierrekorkkinen koeputki (VWR, kataloginnumero 216-0153), jonka sisälle pudotettiin $\varnothing 8$ mm:n teräskuula.

Oireellisista kohdista leikattiin lehtinäytettä steriilillä skalpellilla yhteensä noin 1 cm^2 :n alalta. Oireettomista lehtinäytteistä nypittiin steriileillä pinseteillä kattava otos, noin 15–20 pientä palaa, eri lehdistä. Kummallakin näytteenottomenetelmällä koeputkeen siirretyn näytteen kokonaismäärä oli alle 100 milligrammaa, jotta RNA saadaan puhdistettua riittävän tehokkaasti. Koeputkiin pipetoitiin 500 μl :aa RCL-puskuria, johon oli lisätty ennen käyttöä 2-merkaptotaania (10 μl 2-ME/1 ml puskuria). Koeputkia ravisteltiin kuularavistelijassa kahden minuutin ajan 1000 kierrosta/minuutti. Ravistelu toistettiin, mikäli näytteet eivät olleet jauhautuneet aivan tasaiseksi kasvimehuksi.

Puuvartisten näytteiden jauhaminen tehtiin suodatinpusseissa (Bioreba, Sveitsi) porakoneeseen liitetyllä kuulajauhimella. Kutakin näytettä varten varattiin yksi Bioreban pussi, johon punnittiin 500–600 mg:aa näytettä. Pusseihin lisättiin 5 ml:aa eristyspuskuria (4 M guanidiinitiosyanaatti, 0,2 M natriumasetaatti (pH 5,0), 25 mM EDTA, 2,5 % (paino/tilavuus) PVP-40 ja 1 % 2-merkaptotaani (lisättiin juuri ennen käyttöä)) (MacKenzie ym. 1997). Näytteet jauhettiin tasaiseksi kasvimehuksi.

Jauhetuista näytteistä otettiin lisaattia kertakäyttöisellä pasteur-pipetillä 1,5 ml:n Eppendorf-putkeen 1 ml:n merkkiin asti. Näytteisiin lisättiin 100 μl :aa 20 %:sta sarkosyyliliuosta, minkä jälkeen ne vorteksoitiin kevyesti. Näytteitä inkuboitiin ThermoMixer-lämpöhauteessa (Eppendorf, Saksa) (70°C) ravistellen 1200 kierrosta/minuutti kymmenen minuutin ajan.

RNA:n eristämistä varten kutakin näytettä varten varattiin 2 ml:n keräysputken sisällä oleva QIAshredder spin column -pylväs. Pylväisiin pipetoitiin 450 μl :aa näy-

tettä. Putkia sentrifugoitiin täydellä teholla (21 000 g) kahden minuutin ajan, minkä jälkeen 420 µl:aa supernatanttia siirrettiin 210 µl:aa absoluuttista etanolia sisältävään Eppendorf-putkeen. Supernatantti ja etanoli sekoitettiin toisiinsa välittömästi pipetillä pumpaamalla.

Seos pipetoitiin 2 ml:n keräysputken sisällä olevaan RNeasy spin column -pylväseen. Näytteiden sentrifugoinnin (30 s × 8 000 g) jälkeen pylväs asetettiin uuteen 2 ml:n keräysputkeen. Näytteisiin lisättiin 700 µl:aa RW1-puskuria ja näytteiden sentrifugoinnin (30 s × 8 000 g) jälkeen pylväs asetettiin uuteen 2 ml:n keräysputkeen. Näytteisiin lisättiin 500 µl:aa RPE-puskuria, johon oli lisätty absoluuttista etanolia ennen ensimmäistä käyttökertaa. Näytteiden sentrifugoinnin (30 s × 8 000 g) jälkeen pylväs asetettiin uuteen 2 ml:n keräysputkeen, minkä jälkeen näytteisiin lisättiin uudestaan 500 µl:aa RPE-puskuria. Pylväiden kuivaamiseksi näytteitä sentrifugoitiin 2 min × 8 000 g, minkä jälkeen pylväät asetettiin 1,5 ml:n keräysputkeen. Näytteisiin lisättiin 50 µl:aa RNAasitonta vettä. Näytteitä sentrifugoitiin 1 min × 8 000 g, jolloin puhdistettu RNA kulkeutui pylvästä keräysputken pohjalle. RNA:n konsentraatio ja puhtaus mitattiin NanoDrop ND-1000 -spektrofotometrillä (Thermo Fisher Scientific, Yhdysvallat). RNA-näytteet säilytettiin syväjäähdytyksessä (−80 °C).

4.3 Multiplex-RT-PCR:n optimointi ja suorittaminen

Tutkimuksessa käytettiin kirjallisuudesta etsittyjä yksittäisiä virusheimoja, -sukuja tai alaryhmiä tunnistavia degeneroituja alukkeita (Oligomer Oy, Suomi) (taulukko 2 ja liite A). Kaikki kyseiset alukkeet ovat suunniteltu sitoutumaan tietylle heimolle, suvulle tai alaryhmälle ominaiseen konservoituneeseen alueeseen, jota ei täsmälleen samanlaisena esiinny muilla virusheimoilla, -suvuilla tai suvun muilla alaryhmillä.

Optimointia varten jokaiselle alukeparille oli syväjäähdyksessä (−80 °C) säilytetty RNA-näyte positiivisesta kontrolliviruksesta (taulukko 3). Kontrollivirukset olivat peräisin kaupallisista kantakokoelmista tai Eviran omien tutkimuksien yhteydessä löydettyistä, sekvensoimalla tunnistetuista viruksista. Optimointi suoritettiin testaamalla alukkeet ensin erikseen ja tämän jälkeen yhdessä käyttäen ”SuperScript III One-Step RT-PCR System with Platinum *Taq* DNA polymerase”-pakkausta (kataloginnumero 12574-026, Thermo Fisher Scientific, Yhdysvallat) noudattaen valmistajan ohjeiden mukaisia olosuhteita. Simplex-PCR:ssä yhtä näytettä varten tehty 10 µl:n reaktioseos sisälsi 5 µl:aa 2x reaction mix -reaktioseosta, 1,0 µl:aa forward- ja reverse-alukeseosta (molemmat 10 µM), 0,4 µl:aa SuperScript III RT/Platinum *Taq*

Mix -entsyymiseosta, 0,5 µl:aa kontrolli-RNA:ta ja 3,1 µl:aa steriiliä UP-vettä (ultra pure), joka on puhdistettu käyttäen Milli-Q® Direct Water Purification System -laitteistoa (Merck, Saksa).

RT-PCR:n suorittamista varten syväjäähdytysastimessa (−80 °C) säilytetyt RNA-näytteet ja pakastimessa (−20 °C) säilytetyt alukkeet, steriili DEPC-käsitelty UP-vesi ja SuperScript-pakkauksen sisältämät entsyymiseos ja 2x reaction mix -puskuriliuos sulatettiin juuri ennen käyttöä. Dietyylipyrokarbonaatti-käsittely (DEPC) inaktivoi RNAaseja suojaa näin RNA:ta hajoamiselta. Pipetoinnit tehtiin erillisessä puhdashuoneessa laminaarikaapissa kontaminaatioiden välttämiseksi. Kutakin näytettä ja kontrollia varten varattiin yksi ohutseinäinen 0,2 ml:n PCR-putki tai vaihtoehtoisesti näytteet pipetoitiin PCR-laitteeseen sopivalle 96-kuoppalevyille.

RT-PCR-reaktioseoksiin tarvittava master mix -reaktioliuoksen määrä laskettiin erikseen kullekin reaktioseokselle. Vesi, puskuriliuos ja entsyymiseos sekoitettiin master mix -reaktioliuokseksi, joka vorteksoitiin ja spinnattiin putken pohjalle. Kutakin näytettä sekä positiivista ja negatiivista kontrollia varten pipetoitiin Eppendorf-putkeen 9,5 µl:aa master mix -reaktioliuosta. Kuhunkin Eppendorf-putkeen pipetoitiin 0,5 µl:aa joko näyte- tai kontrolli-RNA:ta ja negatiivisiin kontroleihin vettä.

PCR-ajo suoritettiin 1000 Thermal Cycler (Bio-Rad, Yhdysvallat) tai CFX96 Real-Time PCR System -laitteella (Bio-Rad, Yhdysvallat). PCR-reaktion onnistuminen ja monistuneiden DNA-tuotteiden koko tarkistettiin ajamalla näytteet 1,4 %:ssa TBE-agarosigeelissä (100 V, 1 h 15 min) ja kuvaamalla näytteet UV-valossa. Ennen ajoa 10 µl:n näytteisiin lisättiin 2 µl:aa 20-prosenttista Ficoll-latauspuskuria (Ficoll 10 ml: 2 g Ficoll 400 (Sigma-Aldrich, USA, kataloginnumero F-4375), 2 ml 0,1 M EDTA, pH 8,0, 5 mg bromfenolisini (Merck, Saksa, kataloginnumero 8122) ja 3 mg ksyleenisyanoli (Sigma-Aldrich, USA, kataloginnumero X-4126)) ja laitimmaisiin kaivoihin lisättiin 6 µl:aa 100 bp kokostandardiseosta (New England Biolabs, USA).

Simplex-PCR:ssä yksittäisille alukkeille etsittiin optimaaliset olosuhteet monistamalla positiiviset viruskontrollit eri alukkeiden sitoutumisvaiheen lämpötiloissa ja arvioimalla geelielektroforeesikuvan perusteella sitoutumisvaiheen lämpötilan muutoksen vaikutusta PCR-tuotteiden monistumiseen. Ekstensiovaiheen aika laskettiin Superscript-pakkauksen suositusten mukaan, mikä riippuu monistettavan sekvenssin pituudesta. Käänteiskopiointivaiheelle käytetty lämpötila ja aika oli 48 °C/45 min, alkudenaturaatiovaiheelle 94 °C/2 min ja näitä seuraten 34 kertaa denaturaatiovaihe (95 °C/30 s), alukkeiden sitoutumisvaihe (lämpötilan vaihdel-

lessa välillä 45–60 °C/45 s), ekstensiovaihe (72 °C/1 min/kiloemäs) ja kierrosten lopuksi loppuekstensiovaihe (72 °C/5 min).

Yksittäisten alukkeiden optimoinnista saatuja tuloksia hyödynnettiin multiplex-PCR:n suunnittelussa valitsemalla samaan multiplex-testiin ne alukkeet, joiden optimisitoutumislämpötilat olivat lähellä toisiaan. Simplex-PCR:n tulosten perusteella päädyttiin optimoimaan kaksi erillistä multiplex-PCR-testiä, multiplex A ja multiplex B. Testauksessa kokeiltiin eri alukepariyhdistelmiä ennen kuin löydettiin lopulliset yhdistelmät, joissa kaikki testin alukeparit toimivat. Multiplex-PCR:ssä kutakin näytettä varten tehty reaktioseos ja PCR:ssä käytetty ohjelma tehtiin muokailleen simplex-PCR:ssä tehtyjä reaktioseoksia ja PCR-ohjelmaa. Simplex-PCR:stä poiketen multiplex-PCR:ssä käytettiin neljästä alukeparista koostuvaa seosta, jossa kunkin forward- ja reversealukkeen pitoisuus oli 5 µM. Ekstensiovaiheen aika määräytyi multiplex-PCR:ssä suurimman monistettavan tuotteen mukaan.

Multiplex-PCR:n optimointi suoritettiin testaamalla eri magnesiumsulfaatin pitoisuuksia lisäämällä reaktioseokseen Superscript-pakkauksen mukana erillisessä putkessa tullutta 5 mM magnesiumsulfaattia ja monistamalla positiiviset viruskontrollit eri alukkeiden sitoutumisvaiheen lämpötiloissa (multiplex A: välillä 45–50 °C ja multiplex B: välillä 52–60 °C) ja arvioimalla näiden vaikutusta PCR-tuotteiden monistumiseen. Kokeillut magnesiumsulfaatin pitoisuudet olivat 1,6 mM (ei lisättyä magnesiumsulfaattia), 1,9 mM, 2,1 mM, 2,35 mM ja 2,65 mM.

Optimoinnin jälkeen multiplex-PCR-testien herkkyys arvioitiin suorittamalla testit infektoituneen kasvin kokonais-RNA:sta tehdyillä laimennoksilla, jotka sisälsivät positiivisen kontrolliviruksen. Multiplex A -testin laimennokset tehtiin 1:10, 1:50 ja 1:100 suhteessa RNA-näytettä ja UP-vettä. Multiplex B -testin laimennokset sisälsivät joko 10 ng tai 1 ng infektoituneen kasvin kokonais-RNA:ta. Multiplex-PCR:n optimoinnin perusteella saatu reaktioseos tarkkoine määrineen kutakin näytettä varten ja PCR:ssä käytetyt ohjelmat ovat esiteltynä tuloksissa taulukoissa 4 ja 5.

Käsiteltäessä tutkimukseen aineistoon kuuluvia kasvinäytteitä RT-PCR:n positiivisena sisäisenä kontrollina käytettiin PLANT UNI -kasvialukkeita (PLANT UNI 1 5' – TTT AGT GCT GGT ATG ATC GC – 3' ja PLANT UNI 2 5' – TGG GAA GTC CTC GTG TTG CA – 3') (Kolchinsky ym. 1991). Kyseiset alukkeet tunnistavat eri kasvilajien ribosomaalisen 5S-geenin sekvenssiä. Sekvenssi ilmentyy kaikissa kasvisolukoissa, mistä johtuen PLANT UNI -kasvialukkeita voidaan käyttää

kasvivirusten PCR-testauksessa välttämään vääriä negatiivisia tuloksia (EPPO 2004).

Kasvinäytteiden multiplex-RT-PCR-testauksessa saatu alustava positiivinen tulos varmistettiin tuotteen kokoa vastaavilla alukkeilla (taulukko 2) simplex-reaktiona. Simplex-PCR:n onnistuminen ja monistuneiden DNA-tuotteiden koko tarkistettiin geelielektroforeesin avulla. PCR-testi ei ole viruslajispefinen, mistä johtuen viruslajin määrittämiseksi ja tulosten oikeellisuuden toteamiseksi simplex-reaktiossa saadut positiiviset näytteet lähetettiin suorasekvensoitavaksi Helsingin yliopiston Haartman-instituuttiin. Homologisten sekvenssien löytämiseksi kohdesekvensseille SeqTrace-ohjelmalla (Stucky 2012) koottuja konsensussekvenssejä verrattiin NCBI:n BLASTn- ja BLASTx-työkaluilla tietokannassa oleviin virussekvensseihin käyttäen oletusparametrejä (Altschul ym. 1990).

Taulukko 2. Multiplex-RT-PCR:ssä käytetyt alukkeet, alukkeilla tunnistettavat virusheimot, -suvut tai alaryhmät ja alukkeiden sitoutumiskohdan koodaama proteiinialue. Alukkeiden vaihtelevat emäkset ovat esiteltynä liitteessä A. Tobamoviruksen alaryhmä yhtä monistavat alukeparit voivat tuottaa 685 tai 845 emäsparin pituisia PCR-tuotteita johtuen alukkeen Tob-Uni 1 kahdesta mahdollisesta sitoutumiskohdasta.

Heimo	Suku	Proteiini	5'–3'	3'–5'	Tuote (nt)	Lähde
<i>Alphaflexiviridae</i>	<i>Potexvirus</i>	RdRp	caycarcargcmaargayga	agcatrgcnsrctcytg	584	Vlugt ja Berendsen 2002
<i>Bromoviridae</i>	<i>Alfamovirus, Anulavirus, Bromovirus, Cucumovirus, Ilarvirus, Oleavirus</i>	Metyylitransferaasi	gcnggwtgyggdaarwcnac	amdggwayytgytyngtrtcacc	300	Untiveros ym. 2010
<i>Secoviridae</i>	<i>Nepovirus</i> , alaryhmä A	RdRp	acdtcwgarggitaycc	ratdccyacytgrcwigga	340	Wei ja Clover 2008
<i>Potyviridae</i>	<i>Potyvirus</i>	Inklusioproteiini	gtitgygtigaygaytayaaya	tciaciactigigaigygncc	360	Zheng ym. 2010
<i>Secoviridae</i>	<i>Nepovirus</i> , alaryhmä B	RdRp	tctggittgcyttracrgt	cttrtcactvccatcrgtaa	250	Wei ja Clover 2008
<i>Tombusviridae</i>	<i>Carmovirus, Dianthovirus, Tombusvirus</i>	RdRp	aargtcgacccgwnccnmngntnathcaacc	gmmctgcagnacrcartcrtcnccrtrtt	500	Morozov ym. 1995
<i>Virgaviridae</i>	<i>Tobamovirus</i> , alaryhmä 1	Kuoriproteiini	atttaagtggagggaataaccact	gtygttgatgagttcgtgga	685/845	Letscher ym. 2002
<i>Virgaviridae</i>	<i>Tobravirus</i>	RdRp	ggtggkcaatggtcttwtgg	gtcagctgytgatcagataacc	900	Jones ym. 2008
<i>Bunyaviridae</i>	<i>Tospovirus</i>	RdRp	atgggdatntttgatttcataagrtatgc	tcatgtcatsagrtataatytctc	1200	Chen ym. 2012

Taulukko 3. Multiplex-RT-PCR-menetelmän optimoinnissa käytetyt positiiviset kontrollivirukset.

Suku	Laji	Lyhenne
<i>Potexvirus</i>	Pepinon mosaiikkivirus	PepMV
<i>Ilarvirus</i>	Omenan mosaiikkivirus	ApMV
<i>Nepovirus</i> , alaryhmä a	Viiniköynnöksen viuhkalehtivirus	GFLV
<i>Potyvirus</i>	Luumun rokkovirus	PPV
<i>Nepovirus</i> , alaryhmä b	Tomaatin mustarengaslaikkuvirus	TBRV
<i>Tombusvirus</i>	Tomaatin kitukasvuvirus	TBSV
<i>Tobamovirus</i> , alaryhmä 1	Tupakan mosaiikkivirus	TMV
<i>Tobravirus</i>	Tupakan rattlevirus	TRV
<i>Tospovirus</i>	Palsamin kuoliolaikkuvirus	INSV

4.4 RNA:n eristäminen siRNA-syväsekvensointia varten

RNA eristettiin kasvimateriaalista Chang ym. (1993) ohjetta mukaillen. Pakastimes-
sa (-20°C) säilytettyjä lehtinäytteitä jauhettiin morttelissa nestetypessä ja 100–200
milligrammaa jauhettua lehtinäytettä siirrettiin 2 ml:n Eppendorf-putkeen.

Eristyspuskuri sisälsi 2 % setyylitrimetyyliammoniumbromidia (CTAB), 2 %
polyvinyylipyrrolidonia (PVP, K 30), 100 mM Tris-HCl (pH 8.0), 25 mM EDTA, 2 M
natriumkloridia ja 2 % 2-merkaptetaanolia (lisättiin juuri ennen käyttöä). Näytteisiin
lisättiin 600 μl :aa 65°C :seen esilämmitettyä eristyspuskuriä. Sekoittamisen jälkeen
näytteet siirrettiin lämpöblokkiin (65°C) viiden minuutin ajaksi.

Ylin nestemäinen faasi siirrettiin uuteen 1,5 ml:n Eppendorf-putkeen ja 600 μl :aa
kloroformi:isoamyylialkoholia (24:1) lisättiin näytteisiin. Sekoittamisen jälkeen
näytteitä sentrifugoitiin (15 800 g) viiden minuutin ajan (Eppendorf 5415 R, Ham-
puri, Saksa). Eristäminen kloroformilla toistettiin kahdesti. RNA:n saostamiseksi
ylin nestemäinen faasi siirrettiin uuteen 1,5 ml:n putkeen ja 60 μl 3 M natriumase-
taattia ja 600 μl isopropanolia lisättiin näytteisiin. Näytteitä inkuboitiin 4°C :ssa 18
tunnin ajan.

Näytteet sentrifugoitiin (15 800 g) 30 minuutin ajan 4°C :ssa. Supernatantti poistet-
tiin näytteistä ja pelletit suspensoitiin 200 μl :aan MilliQ-vettä. Näytteisiin lisättiin
100 μl :aa fenolia ja 100 μl :aa kloroformi:isoamyylialkoholia (24:1). Ravistamisen
jälkeen näytteitä sentrifugoitiin (15 800 g) viiden minuutin ajan, minkä jälkeen ylin

nestemäinen faasi siirrettiin uuteen 1,5 ml:n Eppendorf-putkeen. Fenoli-kloroformi-ekstraktio toistettiin kahdesti RNA:n puhtauden lisäämiseksi. RNA saostettiin lisäämällä 200 µl:aa 2 M kaliumasetaattia näytteisiin. Näytteitä inkuboitiin sekoittamisen jälkeen 4 °C:ssa 18 tunnin ajan.

Näytteet sentrifugoitiin (15 800 g) 30 minuutin ajan 4 °C:ssa, ja neste siirrettiin uuteen 1,5 ml:n Eppendorf-putkeen. 40 µl:aa 3 M natriumasetaattia ja 1 ml kylmää absoluuttista etanolia lisättiin näytteisiin. Sekoittamisen jälkeen näytteitä inkuboitiin –20 °C:ssa kahden tunnin ajan. Näytteet sentrifugoitiin (15 800 g) 30 minuutin ajan 4 °C:ssa. Supernatantti poistettiin näytteistä ja pellettejä pestiin kylmällä 80 % etanolilla viiden minuutin ajan sentrifugissa (15 800 g). Etanoli poistettiin ja pellettejä kuivatettiin laminaarissa ilmavirrassa 10 minuutin ajan. Pelletit liuotettiin 40 µl:aan MilliQ-vettä ja säilytettiin –80 °C:ssa.

RNA:n konsentraatio ja puhtaus mitattiin NanoDrop 2000c UV-vis -spektrofotometrillä (Thermo Scientific, USA) valmistajan ohjeiden mukaan. RNA:n konsentraatio mitattiin 260 nm:n aallonpituudella. Absorbanssisuhdetta 260 nm:n ja 280 nm aallonpituudella (A₂₆₀/A₂₈₀) käytetään RNA:n puhtauden selvittämisessä. Puhtaan RNA:n suhdeluku on 1,8–2. RNA:n laatu tarkistettiin visuaalisesti ajamalla RNA-näytteet agarosigeelielektroforeesissa, jossa geeliin oli lisätty etidiumbromidia ja kuvaamalla geelit UV-valossa.

4.5 siRNA-syväsekvensointi

RNA-näytteet lähetettiin kuriirilla kuivajäähän pakattuna syväsekvensointia varten Fasteris SA:hen (Plan-les-Ouates, Sveitsi). Ennen lähettämistä näytteet poolattiin yhteen samaan Eppendorf-putkeen, joka sisälsi 3 µg RNA:ta 10 µl:n tilavuudessa MQ-vettä. Pooli sisälsi saman verran RNA:ta jokaisesta näytteestä.

Syväsekvensoinnissa näytteet ajettiin ensin akryyliamidigeelissä, jonka jälkeen 18–30 nukleotidin pituiset RNA-molekyylit puhdistettiin geelistä. Puhdistamisen jälkeen RNA-molekyyleihin liitettiin 3'-päähän yksijuosteinen adapteri ja 5'-päähän viivakoodattu adapteri. RNA-molekyylit käänteiskopioitiin cDNA:ksi ja monistettiin PCR:llä cDNA-kirjaston luomista varten. Kirjaston tekemisessä käytettiin TruSeq small RNA library -kittiä (Illumina, USA). Kirjaston syväsekvensointi tehtiin Illumina HiSeq 2500 -laitteella (Illumina, USA). Syväsekvensoinnista saatujen tulosten käsittely on laskennallisesti vaativa prosessi, mistä johtuen sekvenssidatan analysoinnissa käytettiin CSC:n laskentapalveluja (CSC - Tieteen tietotekniikan keskus Oy, Espoo)

4.6 Kontigien kokoaminen sRNA-sekvensseistä Velvet-ohjelmalla

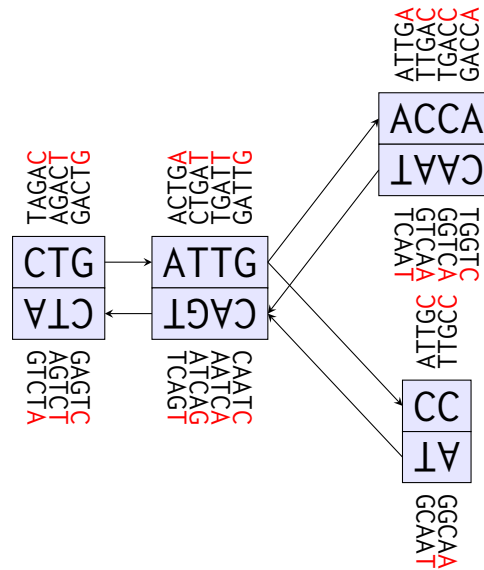
Velvet on ohjelma, jonka avulla lyhyistä (< 50 nt) nukleotidisekvensseistä voidaan koota eripituisia yhtenäisiä sekvenssejä, kontigeja, päällekkäin menevien sekvenssijaksojen perusteella ilman referenssisekvenssiä (Zerbino ja Birney 2008). Velvet-ohjelmaa käytettiin muodostamaan syväsekvensoinnista saaduista sRNA-sekvensseistä (small RNA) kontigeja. DCL-proteiinit tuottavat RNA-hiljennyksessä 21–24 nukleotidin mittaisia siRNA-molekyyliä, mistä johtuen syväsekvensoinnista saadusta aineistosta käytettiin ainoastaan näiden kokoa vastaavia sekvenssejä.

Velvet-ohjelman toiminta perustuu De Bruijn'in graafiin, joka on graafiteorian eräs sovellus. Velvet-ohjelmassa De Bruijn'in graafin muodostaminen tapahtuu katkomalla alkuperäiset sekvenssit osiin, joiden pituus on k . Tätä k -pituisten sekvenssipätkien joukkoa kutsutaan nimellä k -mer. Kaksi eri k -pituista sekvenssiä yhdistetään graafissa toisiinsa, mikäli ensimmäisen sekvenssin viimeiset $k - 1$ nukleotidia vastaa toisen sekvenssin ensimmäistä $k - 1$ nukleotidia (kuva 3). Graafin muodostaminen perustuu siis alkuperäisten sekvenssien sijaan k -mereihin (Zerbino ja Birney 2008). Graafin kokoamisen jälkeen ohjelma yksinkertaistaa graafia ja korjaa siinä esiintyviä virheitä.

Velvet-ohjelman tärkein parametri, k -mer-arvo, vaikuttaa De Bruijn'in graafin rakentumiseen ja siten ohjelman tulosten herkkyyteen ja valikoituvuuteen (Zerbino 2010, Zerbino ja Birney 2008). Pieni k -mer-arvo kasvattaa tulosten herkkyyttä, mutta samalla vähentää niiden valikoivuutta. Suuri k -mer-arvo vaikuttaa tuloksiin päinvastaisesti (Zerbino 2010, Zerbino ja Birney 2008).

K -mer-arvon on oltava pariton luku, jotta se ei olisi käänteisesti komplementaarinen omalle sekvenssilleen. Parilliset k -merit voivat aiheuttaa vääriä toistoja koottaviin sekvensseihin. K -merin on oltava kooltaan pienempi kuin alkuperäiset sekvenssit, sillä k -merejä lyhyempiä sekvenssejä ei huomioida kontigien kokoamisessa. Optimi k -mer-arvo riippuu genomista, alkuperäisten sekvenssien pituudesta ja sekvenssikirjaston laadusta ja kattavuudesta. Velvet-ohjelman kehittäjän mukaan optimi k -mer-arvo on yleensä todettu olevan lähellä 15:tä (Zerbino ja Birney 2008).

K -mer-arvon lisäksi Velvet-ohjelmassa on muita säädettäviä parametreja, jotka vaikuttavat eri tavoin ohjelman tuloksiin. Parametrien optimoinnin avuksi on olemassa eri ohjelmia kuten VelvetOptimiser (Gladman ja Seemann 2012) tai KmerGenie (Chikhi ja Medvedev 2014). Parametrien optimointi voidaan tehdä myös kokeilemalla ohjelmassa eri parametriarvoja ja arvioimalla näiden vaiku-



Kuva 3. Velvet-ohjelmassa De Bruijn'in graafi koostuu k-mereistä ja yhteyksistä, jotka syntyvät toisiaan muistuttavien k-merien välille (Zerbino ja Birney 2008). Kuvan laatikoiden ylä- ja alapuolella esiintyvät sekvenssit esittävät k-merejä ($k = 5$), joista edeltävän k-merin viimeiset $k - 1$ nukleotidia vastaa seuraavan k-merin ensimmäistä $k - 1$ nukleotidia. Koska DNA on kaksijuosteista, rakentuu graafi myös k-mereille käänteisesti komplementaarista k-mereistä, jotka esiintyvät kuvan laatikoiden vastakkaisilla puolilla. Laatikoiden sisällä esiintyvät kirjaimet esittävät kunkin k-merin viimeisintä nukleotidia, joiden perusteella De Bruijn'in graafi luodaan.

tusta tuloksiin (Zerbino ja Birney 2008), mitä menetelmää käytettiin tässä työssä. Velvet-ohjelma ajettiin k-mer-arvoilla 13, 15 ja 17 ja "pienin käytetty kontigin pituus" (minimum contig length) arvoilla 60, 80 ja 100. Jälkimmäinen parametri suodattaa aineistosta pois arvoa lyhyemmät kontigit (Zerbino ja Birney 2008). "Kattavuusrajan" (coverage cutoff) arvona käytettiin viittä kaikissa ajoissa. Parametri suodattaa aineistosta pois sekvenssejä, joissa samaan kohtaan tulevien samanlaisten sekvenssipätkien määrä jää alhaiseksi (Zerbino ja Birney 2008).

Ohjelmassa muodostettuja kontigeja verrattiin NCBI:n BLAST-työkalulla tietokannassa oleviin virussekvensseihin (Altschul ym. 1990). BLAST-hauissa käytettiin oletusparametrejä. Saatua BLAST-osumia käytettiin Bowtie 2 -ohjelmassa referenssisekvensseinä selvittämään, kuinka samanlaisia näytteiden virukset olivat verrattuna tietokantojen viruksiin.

4.7 Kontigien kokoaminen sRNA-sekvensseistä VirusDetect-ohjelmalla

Edellä kuvatun analyysin lisäksi tutkielmassa käytettiin Zhengin ym. (2017) kehittämää, useita analyysijä yhdistelevää VirusDetect-ohjelmaa selvittämään kasvinäytteissä esiintyvät virukset. Ohjelmien tapa koota kontigeja eroaa toisistaan ja siksi ohjelmien antamat tulokset voivat teoriassa poiketa toisistaan.

VirusDetect kokoaa sRNA-sekvensseistä kontigeja kahdella eri tapaa linjaamalla ensin sRNA-sekvenssejä GenBank-tietokannan virussekvensseihin BWA-ohjelmalla (Burrows-Wheeler Alignment) (Li ja Durbin 2009). Pällekkäiset sRNA-sekvenssit yhdistetään toisiinsa SAMtools-ohjelmalla (Li ym. 2009). Tämän jälkeen ohjelma kokoaa kontigeja ilman referenssisekvenssejä Velvet-ohjelman avulla. VirusDetect suorittaa Velvetin parametreille automaattisen optimoinnin. Kummallakin menetelmällä kootut kontigit yhdistetään samaksi aineistoksi, jota verrataan NCBI:n tietokannan virusten nukleotidisekvensseihin BLASTn-työkalulla (Altschul ym. 1990). Jäljelle jääneitä kontigeja, joille ei löydetty vastaavuuksia BLASTn-haussa, verrataan virusten proteiinisekvensseihin BLASTx-työkalulla (Altschul ym. 1990).

4.8 sRNA-sekvenssien linjaaminen referenssisekvensseihin Bowtie 2 -ohjelmalla

Bowtie 2 -ohjelmalla voidaan linjata lyhyitä DNA-sekvenssejä referenssisekvensseihin (Langmead ja Salzberg 2012). Tutkielmassa Bowtie 2:ta käytettiin linjaamaan sRNA-sekvenssejä Velvet- ja Virusdetect-ohjelmien avulla löydettyihin referenssisekvensseihin. Linjaamista käytettiin selvittämään kuinka ison osan ja kuinka monta kertaa keskimäärin sRNA-sekvenssit peittävät referenssigenomia. Linjaamisen valittiin referenssivirusten kokopitkä sekvenssi.

Ennen linjaamista referenssisekvenssit indeksoitiin erityisellä algoritmilla (Ferragina ja Manzini 2000), mikä nopeutti linjaamisen tekoa. Ohjelmassa on useita säädettäviä parametreja, jotka vaikuttavat linjaamisen suorittamiseen ja sen pisteyttämiseen (Langmead ym. 2017). Ohjelmassa käytettiin oletusparametreja paitsi yhden parametrin osalta (parametri -N), joka sai arvoksi nolla ja yksi. Kyseinen parametri vaikuttaa siihen, kuinka monen nukleotidin ero linjauksessa sallitaan. Mitä enemmän nukleotidien eroja linjauksessa sallitaan, sitä enemmän tulosten herkkyys kasvaa ja valikoivuus pienenee.

Sekvenssien linjaamiseen on olemassa myös muita ohjelmia, kuten MAQ (Li ym. 2008) tai Novoalign (Novocraft Technologies Sdn Bhd. 2014), mutta ohjelmien

käyttämät algoritmit eroavat toisistaan ja siten eri ohjelmat voivat tuottaa erilaisia tuloksia samalla datalla. Bowtie 2:ta saatujen tulosten visualisointiin käytettiin Tablet-ohjelmaa (Milne ym. 2013).

4.9 Sekvenssien fylogeneettinen analyysi

Fylogeneettisen analyysin avulla tarkasteltiin multiplex-RT-PCR:stä saatujen virus-sekvenssien sukulaisuussuhteita saman lajin muihin isolaatteihin tai saman suvun muihin viruslajeihin. Usean sekvenssin rinnastukset tehtiin ClustalW-algoritmilla (Larkin ym. 2007). Sekvenssirinnastuksista muodostettiin fylogeneettisiä puita MEGA X ohjelmalla (Kumar ym. 2018) "Neighbor-Joining"-menetelmällä. Menetelmässä muodostetaan sekvensseistä ryhmiä pyrkien löytämään etäisyyspuu, jonka oksien kokonaispituus saa pienimmän mahdollisen arvon (Saitou and Nei 1987). Puiden luotettavuus arvioitiin bootstrap-menetelmällä (Efron 1979), jossa arvot testattiin 1000:n toiston mukaan. Puumallien evolutiivisia etäisyyksiä laskettiin Kimura-2-parametrin avulla (Kimura 1980). Clustal Omega -ohjelmaa käytettiin laskemaan nukleotidiyhtäläisyyksiä näytteistä löydettyjen virussekvenssien ja referenssisekvenssien välille (McWilliam ym. 2013).

5 Tulokset

5.1 Multiplex-RT-PCR:n optimoinnin tulokset

Simplex-PCR:n alukkeiden sitoutumislämpötilan optimoinnin jälkeen kaikki virusalukkeet monistivat PCR:ssä odotetun kokoisia, selkeitä tuotteita ilman epäspesifistä monistumista. Simplex-PCR:n sitoutumisvaiheen lämpötilan optimointitulosten perusteella sekä eri alukepariyhdistelmien kokeilun jälkeen multiplex A -testiin valitiin nepovirusten alaryhmää a, *Bromoviridae*-heimoa sekä potex- ja potyvirussukua tunnistavat alukeparit. Multiplex B -testiin valitiin nepovirusten alaryhmää b, tobnavirussukua, tobamovirus alaryhmä yhtä ja *Tombusviridae*-heimoon kuuluvien carmo-, diantho- ja tombusvirussukuja tunnistavat alukkeet. Tospovirusia tunnistavat alukkeet rajattiin pois tutkimuksesta, sillä ne eivät tuottaneet multiplex-PCR:ssä eri alukepariyhdistelmillä näkyvää PCR-tuotetta palsamin kuoliolaikkuviruksesta (INSV) tai muodostunut tuote ei vastannut odotetun kokoista tuotetta (1200 nt).

Magnesiumsulfaatin konsentraation kasvattaminen vaikutti positiivisesti sekä multiplex A -testin että multiplex B -testin herkkyyteen havaita positiivisia kontrolliviruksia. Kummassakin testissä positiivisista kontrolliviruksista monistuneet PCR-tuotteet olivat parhaiten havaittavissa Mg^{2+} -konsentraation ollessa 2,65 mM. Kokeiltujen alukkeiden sitoutumislämpötilojen perusteella multiplex A:ssa positiiviset viruskontrollit monistuivat tehokkaimmin alukkeen sitoutumislämpötilan ollessa 48 °C ja multiplex B:ssä alukkeen sitoutumislämpötilan ollessa 57 °C (taulukot 4 ja 5).

Multiplex A -testin herkkyytulosten perusteella (kuva 4 ja taulukko 6) alhaisin havaittu pitoisuus infektoituneen kasvin kokonais-RNA:sta oli viiniköynnöksen viuhkalehtivirusta (GFLV) sisältävien laimennosten suhteen 4,5 ng/μl, omenan mosaiikkivirusta (ApMV) sisältävien laimennosten suhteen 2,5 ng/μl, luumun rokkovirusta (PPV) sisältävien laimennosten suhteen 0,1 ng/μl ja pepinon mosaiikkivirusta (PepMV) sisältävien laimennosten suhteen 1,8 ng/μl. Geelielektroforeesikuvasta on havaittavissa oikean kokoisten tuotteiden lisäksi epäspesifisiä tuotteita.

Multiplex B -testin herkkyytulokset (kuva 5 ja taulukko 6) osoittavat testissä alhaimman havaitun ja testatun pitoisuuden infektoituneen kasvin kokonais-RNA:sta olleen 1 ng/μl kaikista eri positiivisia viruskontrolleja sisältäneistä laimennoksista.

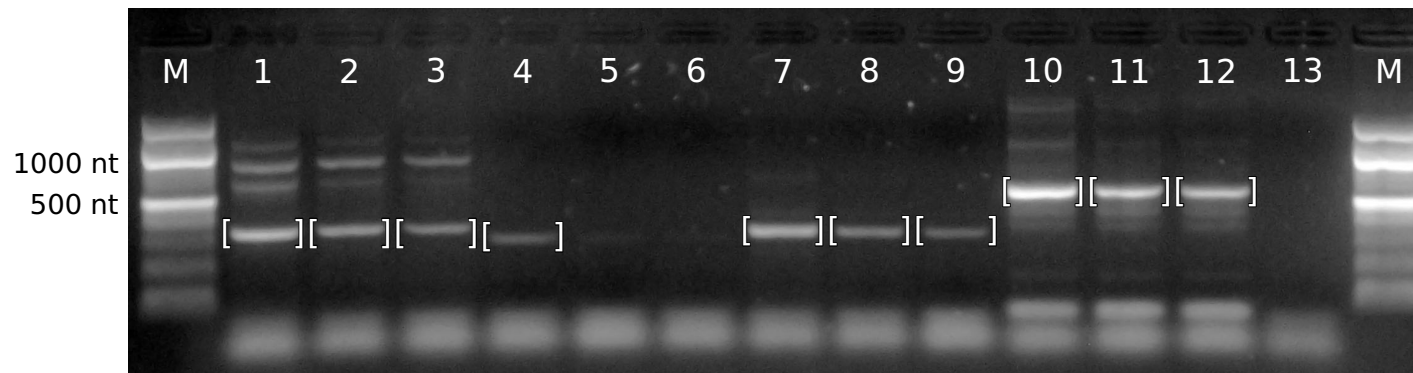
Geelielektroforeesikuvasta on havaittavissa oikean kokoisten tuotteiden lisäksi epäspesifisiä tuotteita sekä kohdesekvenssien epätasaista monistumista.

Taulukko 4. Reaktioseos (10 µl) yhtä putkea varten ja PCR-ohjelma alukeseos multiplex A:lle.

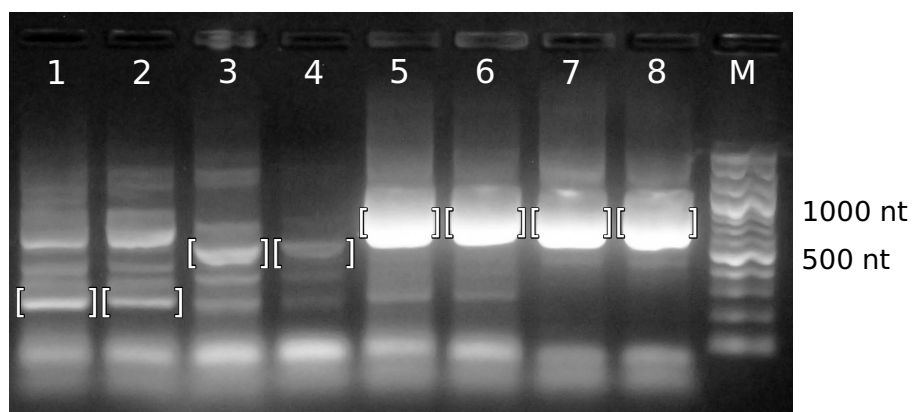
Reagenssi		PCR-sykli		
H ₂ O	1 µl	1×	45 min	48 °C
2x Reaction Mix	5 µl	1×	2 min	94 °C
5 mM Mg	2,1 µl	34×	30 s	95 °C
Alukeseos A 5 µM	1 µl		45 s	48 °C
SuperScript III RT/Platinum <i>Taq</i> Mix	0,4 µl		1 min 20 s	72 °C
RNA näyte	0,5 µl	1×	5 min	72 °C
Yhteensä	10 µl			

Taulukko 5. Reaktioseos (10 µl) yhtä putkea varten ja PCR-ohjelma alukeseos multiplex B:lle.

Reagenssi		PCR-sykli		
H ₂ O	1 µl	1×	45 min	48 °C
2x Reaction Mix	5 µl	1×	2 min	94 °C
5 mM Mg	2,1 µl	34×	30 s	95 °C
Alukeseos B 5 µM	1 µl		45 s	57 °C
SuperScript III RT/Platinum <i>Taq</i> Mix	0,4 µl		1 min 20 s	72 °C
RNA näyte	0,5 µl	1×	5 min	72 °C
Yhteensä	10 µl			



Kuva 4. Geelielektroforeesikuva multiplex A -testin herkkyystuloksista. PCR-reaktiot sisältävät infektoituneen kasvin kokonais-RNA:sta tehtyjä laimennoksia, jotka sisälsivät positiivisen kontrolliviruksen. Näytteiden järjestys: 1: viiniköynnöksen viuhkalehtivirus (GFLV) (1:10), 2: GFLV (1:50), 3: GFLV (1:100), 4: omenan mosaiikkivirus ApMV (1:10), 5: ApMV (1:50), 6: ApMV (1:100), 7: luumun rokkovirus PPV (1:10), 8: PPV (1:50), 9: PPV (1:100) 10: pepinon mosaiikkivirus PepMV (1:10), 11: PepMV (1:50), 12: PepMV (1:100) 13: negatiivinen vesikontrolli. Oikean kokoiset PCR-tuotteet ovat merkitty erottamaan ne epäspesifisistä tuotteista.



Kuva 5. Geelielektroforeesikuva multiplex B -testin herkkyystuloksista. PCR-reaktiot sisältävät infektoituneen kasvin kokonais-RNA:sta tehtyjä laimennoksia, jotka sisälsivät positiivisen kontrolliviruksen. Näytteiden järjestys: 1: tomaatin mustarengaslaikkuvirus TBRV (10 ng), 2: TBRV (1 ng), 3: tomaatin kitukasvuvirus TBSV (10 ng), 4: TBSV (1 ng), 5: tupakan rattlevirus TRV (10 ng), 6: TRV (1 ng), 7: tupakan mosaiikkivirus TMV (10 ng) ja 8: TMV (1 ng). Oikean kokoiset PCR-tuotteet ovat merkitty erottamaan ne epäspesifisistä tuotteista.

Taulukko 6. Multiplex-PCR-testeillä alin havaittu ja alin testattu pitoisuus infektoituneen kasvin kokonais-RNA:sta, joka sisälsi positiivisen kontrolliviruksen.

Multiplex-testi	Alukkeiden tunnistama virusheimo, -suku tai suvun alaryhmä	Positiivinen kontrollivirus	Laimentamattoman RNA:n pitoisuus (ng/μl)	Alin havaittu pitoisuus (ng/μl)	Alin testattu pitoisuus (ng/μl)
A	<i>Nepovirus</i> , alaryhmä a	Viiniköynnöksen viuhka-lehtivirus, GFLV	454,8	4,5	4,5
A	<i>Bromoviridae</i>	Omenan mosaiikkivirus, ApMV	25	2,5	0,25
A	<i>Potyvirus</i>	Luumun rokkovirus, PPV	13,9	0,1	0,1
A	<i>Potexvirus</i>	Pepinon mosaiikkivirus, PepMV	179,5	1,8	1,8
B	<i>Nepovirus</i> , alaryhmä b	Tomaatin mustarengas-laikkuvirus, TBRV	138,6	1	1
B	<i>Carmovirus</i> , <i>Dianthovirus</i> , <i>Tombusvirus</i>	Tomaatin kitukasvuvirus, TBSV	827	1	1
B	<i>Tobravirus</i>	Tupakan rattlevirus, TRV	15,2	1	1
B	<i>Tobamovirus</i> , alaryhmä 1	Tupakan mosaiikkivirus, TMV	234,9	1	1

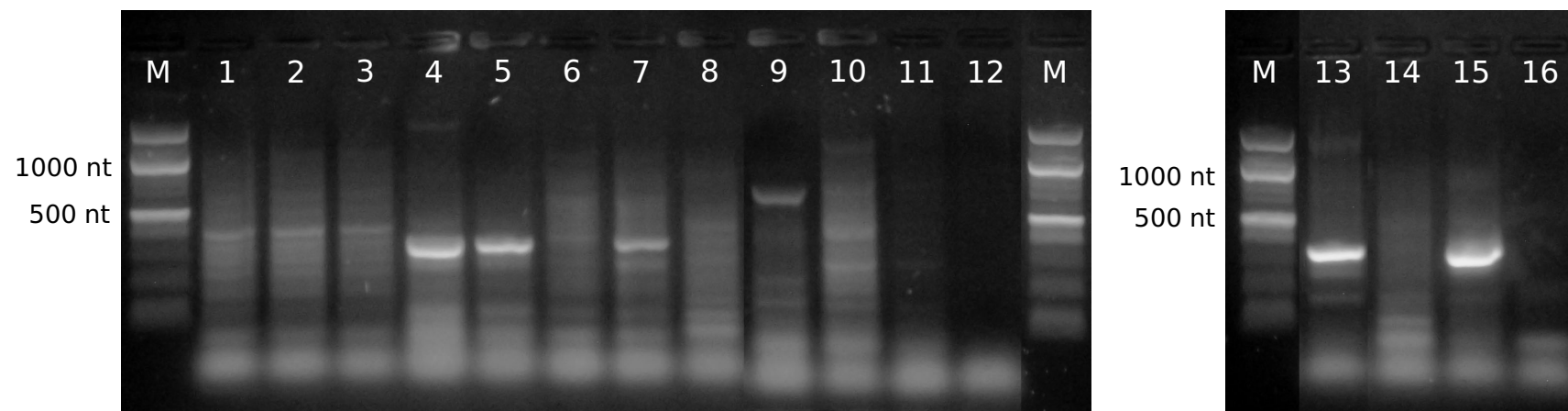
5.2 Multiplex-RT-PCR:llä tunnistetut kasvivirukset

Multiplex A -testissä näytteistä 787 (kiinanjaloangervo), 788 (kiinanjaloangervo), 790 (kvinoa), JT12 (lupiini) ja JT15 (valkomesikkä) muodostui noin 350 emäsparin kokoinen tuote (kuva 6). Kasvinäytteiden multiplex-RT-PCR-testauksessa saatu alustava positiivinen tulos varmistettiin tuotteen kokoa vastaavilla alukkeilla (taulukko 2) simplex-reaktiona. Tuotteiden kokoa vastaavat alukkeet olivat potyvirusia ja nepovirusten alaryhmä a:ta tunnistavat alukkeet. Simplex-PCR-testissä kiinanjaloangervo- ja kvinoanäytteistä muodostui noin 350 emäsparin kokoinen tuote käytettäessä nepovirusten alaryhmä a:ta tunnistavia alukkeita ja lupiini- ja valkomesikkänäytteiden kohdalla käytettäessä potyviruskua tunnistavia alukkeita. Näytteestä JT2 (suikeroalpi) muodostui haaleana näkyvä, noin 800 emäsparin kokoinen PCR-tuote. Tämä ei vastannut multiplex A -testissä käytettävien alukkeiden monistamien tuotteiden kokoa, mistä johtuen tulos katsottiin negatiiviseksi.

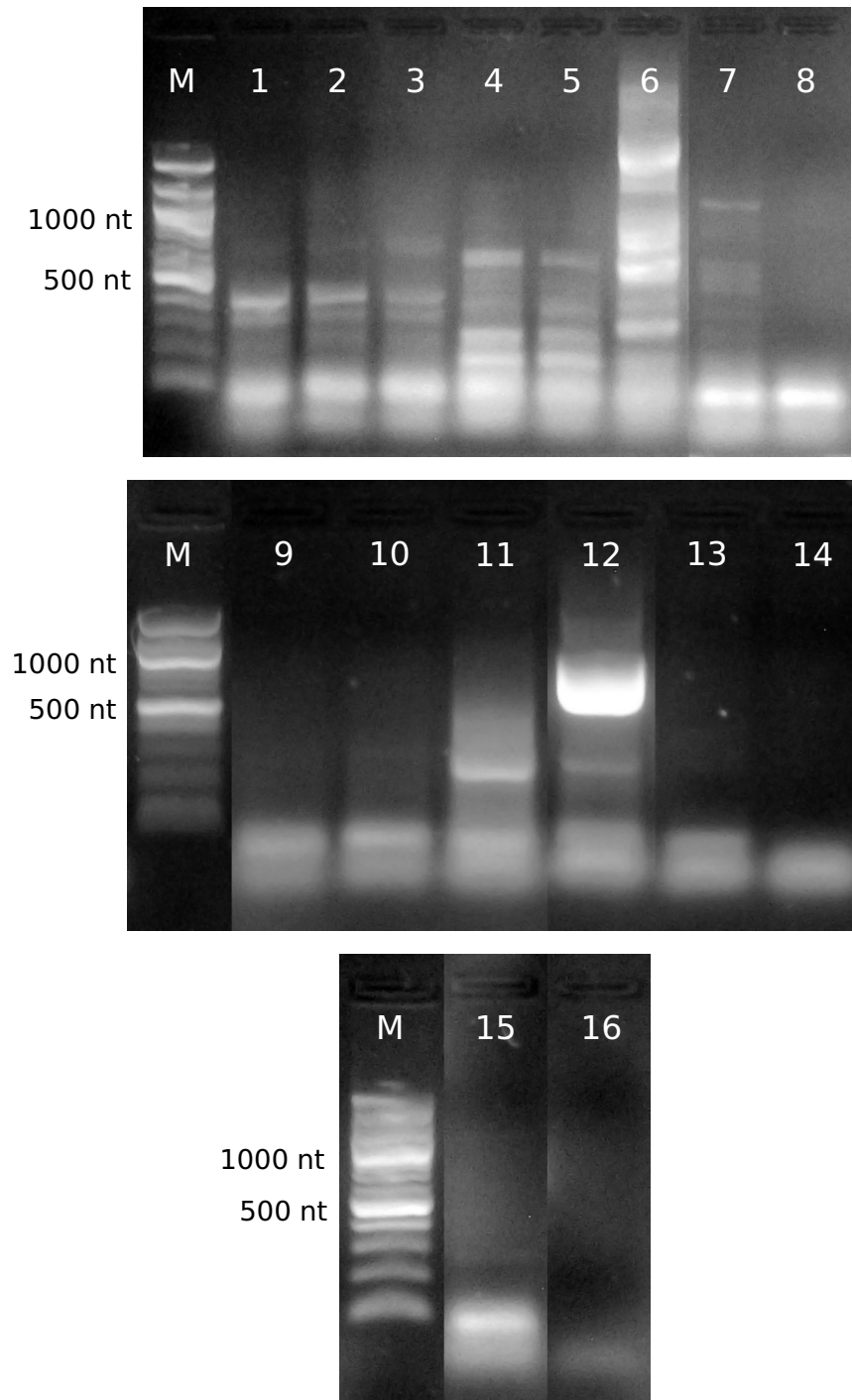
Multiplex B -testissä näytteestä JT2 (suikeroalpi) muodostui noin 900 emäsparin kokoinen tuote (kuva 7). Tuotteen kokoa vastaavat alukkeet olivat tobnavirusia ja tobnavirusten alaryhmä yhtä tunnistavat alukkeet. Simplex-PCR-testissä suikeroalpinäytteestä muodostui 900 emäsparin kokoinen käytettäessä tobnavirusia tunnistavia alukkeita. Näytteestä 791 (alppiruusu) muodostui noin 250 emäsparin kokoinen tuote, mikä vastasi nepovirusten alaryhmä b:tä tunnistavia alukkeita. Alppiruusunäytteestä ei muodostunut tuotetta simplex-PCR-testissä käytettäessä nepovirusten alaryhmä b:tä tunnistavia alukkeita, mistä johtuen tulos katsottiin negatiiviseksi. Osasta kasvinäytteistä muodostui haaleasti näkyviä PCR-tuotteita, mistä johtuen multiplex B -testi uusittiin. Testin uusinnassa kasvinäytteistä ei ilmennyt uusia, selkeästi havaittavissa ja epäspesifisistä tuotteista erotettavissa olevia oikean kokoisia PCR-tuotteita.

Sekvensoiduista PCR-näytteistä tehdyt BLAST-tietokantavertailut osoittivat kasvinäytteistä 787 (kiinanjaloangervo), 788 (kiinanjaloangervo), 790 (kvinoa), JT2 (suikeroalpi) ja JT12 (lupiini) monistuneiden virussekvenssien olevan merkittävästi samankaltaisia tietokannan virussekvenssien kanssa (taulukot 7 ja 8 ja liite B). Kasvinäytteistä 787, 788 ja 790 löytyneet virussekvenssit vastasivat nukleotidisallalla 91-prosenttisesti ja proteiinitasolla 98–99-prosenttisesti NCBI:n tietokantaan tallennettua arabiksen mosaiikkivirussekvenssiä (ArMV). Näytteestä JT2 löydetty virussekvenssi oli huomattavan samankaltainen NCBI:n tietokantaan tallennetun tupakan rattlevirussekvenssin (TRV) kanssa, sillä sekä nukleotidi- että proteiinitasolla identtisyysprosentti oli 99 % ja E-arvo 0,0. Näytteen JT12 virussekvenssi vas-

tasi nukleotiditasolla 98-prosenttisesti ja proteiinitasolla 100-prosenttisesti NCBI:n tietokantaan tallennettua lupiinin mosaiikkivirussekvenssiä (LuMV). Näytteen JT15 (valkomesikkä) virussekvenssi ei antanut täsmällisiä osumia tietokantasekvensseihin, mutta kaikki löydetyt vastaavuudet olivat potyvirussukuun kuuluvia lajeja.



Kuva 6. Multiplex A -testin tulokset testatuista kasvinäytteistä. Näytteiden järjestys geelillä: 1: 870 (kuunlilja), 2: 871 (kuunlilja), 3: 872 (kuunlilja), 4: 787 (kiinanjaloangervo), 5: 788 (kiinanjaloangervo), 6: 789 (kvinoa), 7: 790 (kvinoa), 8: 791 (alppiruusu), 9: JT2 (suikeroalpi), 10: JT3 (keltamo), 11: JT4 (valvatti), 12: JT5 (poimulehti), 13: JT12 (lupiini) 14: JT17 (nokkonen), 15: JT15 (valkomesikkä) ja 16: negatiivinen vesikontrolli.



Kuva 7. Multiplex B -testin tulokset testatuista kasvinäytteistä. Näytteiden järjestys geelikuvissa: 1: 870 (kuunlilja), 2: 871 (kuunlilja), 3: 872 (kuunlilja), 4: 787 (kiinanjalangervo), 5: 788 (kiinanjalangervo), 6: JT3 (keltamo), 7: JT15 (valkomesikkä), 8: negatiivinen vesikontrolli, 9: 789 (kvinoa), 10: 790 (kvinoa), 11: 791 (alppiruusu), 12: JT2 (suikeroalpi), 13: JT4 (valvatti), 14: JT5 (poimulehti), 15: JT12 (lupiini) ja 16: JT17 (nokkonen).

Taulukko 7. BLASTn-haun tulokset multiplex-RT-PCR:llä löydetystä virussekvensseistä. Taulukkoon on valittu se referenssisekvenssi, johon näytteestä havaittu sekvenssi parhaiten linjautui.

Koodi	Suomenkielinen nimi	Virallinen nimi	Lyhenne	Sekvenssin yksilöllinen tunnistus NCBI:ssä	Linjautumisen pituus nukleotideina	Identtisyys (%)	E-arvo
787	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	EU741688.1	285	91	2×10^{-102}
788	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	EU741688.1	310	91	2×10^{-114}
790	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	EU741688.1	320	91	7×10^{-118}
JT2	Tupakan rattlevirus	<i>Tobacco rattle virus</i>	TRV	JX144383.1	721	99	0
JT12	Lupiinin mosaiikkivirus	<i>Lupinus mosaic virus</i>	LuMV	EU847625.2	314	98	4×10^{-155}
JT15	Japanin jamssin mosaiikkivirus	<i>Japanese yam mosaic virus</i>	JYMV	KJ789141.1	269	80	2×10^{-47}

Taulukko 8. BLASTx-haun tulokset multiplex-RT-PCR:llä löydetystä virussekvensseistä. Taulukkoon on valittu se referenssisekvenssi, johon näytteestä havaittu sekvenssi parhaiten linjautui.

Koodi	Suomenkielinen nimi	Virallinen nimi	Lyhenne	Sekvenssin yksilöllinen tunnistus NCBI:ssä	Linjautumisen pituus aminohappoina	Identtisyys (%)	E-arvo
787	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	ACF77019.1	102	99	4×10^{-66}
788	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	ACF77019.1	103	99	4×10^{-67}
790	Arabiksen mosaiikkivirus	<i>Arabis mosaic virus</i>	ArMV	ACF77019.1	106	98	1×10^{-68}
JT2	Tupakan rattlevirus	<i>Tobacco rattle virus</i>	TRV	ALG76062.1	240	99	0
JT12	Lupiinin mosaiikkivirus	<i>Lupinus mosaic virus</i>	LuMV	YP_004123939.1	105	100	7×10^{-72}
JT15	Pavun keltamosaiikkivirus	<i>Bean yellow mosaic virus</i>	BYMV	AEH43856.1	97	93	7×10^{-63}

5.3 siRNA-syväsekvensoinnin tulokset

Syväsekvensoinnista saatiin sRNA-sekvenssejä yhteensä 17 539 516 kappaletta, joista 26,9 % oli pituudeltaan 21–24 nukleotidin mittaisia. Aineistossa 21 nukleotidin mittaisia sekvenssejä oli 12,1 %, 22 nukleotidin mittaisia 5,3 %, 23 nukleotidin mittaisia 1,8 % ja 24 nukleotidin mittaisia 7,8 %. Vajaa puolet (46,2 %) sekvensseistä oli pituudeltaan 1–17 nukleotidin mittaisia. Illumina-sekvensointilaite määritteli poolatulle näytteelle keskimääräiseksi laatuarvoksi 35,59, mikä vastaa 0,028 %:n virhetodennäköisyyttä. Sekvensseistä yli 95 % ylitti laatuarvon 30 (0,1 %:n virhetodennäköisyys).

5.4 Velvet-ohjelmalla tunnistetut virukset

Kokeiltujen parametriarvojen perusteella kontigien kokonaismäärä ja yhteenlaskettu pituus oli suurin käytettäessä k-mer-arvoa 15 ja ”pienin käytetty kontigin pituus”-arvoa 60 (taulukko 9). Kontigien keskimääräinen pituus ja N50-arvo oli suurin käytettäessä k-mer-arvoa 15 ja ”pienin käytetty kontigin pituus”-arvoa 100. N50-arvo tarkoittaa, että 50 % kaikista nukleotideista on peräisin kontigeista, joiden pituus on vähintään N50-arvon verran. Virussekvensseihin osuneiden BLAST-osumien määrä oli suurin käytettäessä k-mer-arvoa 17 ja ”pienin käytetty kontigin pituus”-arvoa 60. Pisimmät kontigit saatiin käytettäessä ohjelmassa k-mer-arvoa 15. BLAST-haun tuloksia tarkasteltiin tuloksista, jotka saatiin parametriarvoilla k-mer 15 ja ”pienin käytetty kontigin pituus”-arvolla 80 (taulukko 10), koska kyseiset parametriarvot sijoittuvat niiden parametriarvojen väliin, joilla saatiin suurin kontigien keskimääräinen pituus ja suurin kontigien kokonaismäärä.

Moni Velvet-ohjelman kontigeista vastasi BLASTn-haun perusteella NCBI:n tietokantaan tallennettuja arabiksen mosaiikkiviruksen (ArMV), tupakan rattleviruksen (TRV) ja lupiinin mosaiikkiviruksen (LuMV) sekvenssejä. Kustakin viruslajista tarkasteltiin alhaisimman E-arvon antanutta linjausta. Tarkastellut kontigit osoittivat merkittävää samankaltaisuutta ArMV:n, TRV:n ja LuMV:n kanssa, identtisyyden vaihdellessa 87 ja 100 prosentin välillä ja E-arvon ollessa 0. Kontigeille löydettiin osittaisia vastaavuuksia nepoviruksen alaryhmä a:han luokiteltavaan viiniköynnöksen viuhkalehtivirukseen (GFLV) ja eri potyvirusiin (pavun keltamosaiikkivirus (BYMV) identtisyys 80,7 %, E-arvo 6×10^{-98} , soijapavun mosaiikkivirus (SMV) identtisyys 79,3 %, E-arvo 1×10^{-17} ja apilan keltasuonivirus (CIYVV) identtisyys 88,4 %, E-arvo 4×10^{-24}).

Taulukko 9. Velvet-ohjelman kokoamien kontigien kokonaismäärä, yhteenlaskettu ja keskimääräinen pituus, N50-arvo ja pisin kontigi nukleotideina sekä BLASTn-haun perusteella virussekvensseihin linjautuneiden kontigien määrä eri parametriarvoilla. N50-arvo tarkoittaa, että 50 % kaikista nukleotideista on peräisin kontigeista, joiden pituus on vähintään N50-arvon verran.

K-mer	"Pienin käytetty kontigin pituus"-arvo	Kokonaismäärä	Yhteenlaskettu pituus	Keskimääräinen pituus	N50-arvo	BLAST-osumien määrä	Pisin kontigi
13	60	246	18 517	75	73	95	152
13	80	67	6 410	96	92	35	152
13	100	19	2 215	117	111	11	152
15	60	878	100 425	114	114	198	1 034
15	80	457	71 742	157	171	146	1 034
15	100	276	55 697	202	210	119	1 034
17	60	692	73 044	106	105	225	782
17	80	349	49 676	142	140	125	782
17	100	224	38 613	172	174	114	782

Taulukko 10. Velvet-ohjelmassa koottujen kontigien linjautuminen NCBI:n tietokannan virussekvensseihin. Taulukossa esitellään viruslajeihin linjautuneiden kontigien lukumäärä ja linjautumisen pituutta ja laatua kuvaavat tulokset alhaisimman E-arvon antaneesta linjauksesta. Tulokset on saatu käyttäen Velvet-ohjelmassa parametriarvoja k-mer 15 ja ”pienin käytetty kontigin pituus”-arvoa 80.

Virallinen nimi	Suomenkielinen nimi	Lyhenne	Suku	Sekvenssin yksilöllinen tunniste NCBI:ssä	Referenssi-sekvenssin pituus (nt)	Linjautumisen pituus	Identisyys (%)	E-arvo	Kontigien määrä
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikki-virus	ArMV	<i>Nepovirus</i>	EU617326 (RNA1)	7 334	618	88,5	0	16
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikki-virus	ArMV	<i>Nepovirus</i>	EU617327 (RNA2)	3 812	772	87,2	0	9
<i>Tobacco rattle virus</i>	Tupakan rattlevirus	TRV	<i>Tobravirus</i>	GQ903771 (RNA1)	6 791	406	99,5	0	29
<i>Tobacco rattle virus</i>	Tupakan rattlevirus	TRV	<i>Tobravirus</i>	KJ826366 (RNA2)	4 154	411	98,3	0	7
<i>Lupinus mosaic virus</i>	Lupiinin mosaiikki-virus	LuMV	<i>Potyvirus</i>	EU847625	10 113	723	98,2	0	38
<i>Bean yellow mosaic virus</i>	Pavun keltamosaiikki-virus	BYMV	<i>Potyvirus</i>	HG970867	9 417	466	80,7	6×10^{-98}	4
<i>Soybean mosaic virus</i>	Soijapavun mosaiikki-virus	SMV	<i>Potyvirus</i>	KP710868	9 994	135	79,3	1×10^{-17}	2
<i>Clover yellow vein virus</i>	Apilan keltasuoni-virus	CIYVV	<i>Potyvirus</i>	HG970870	9 439	96	87,5	4×10^{-22}	2
<i>Grapevine fanleaf virus</i>	Viiniköynnöksen viuhkalehtivirus	GFLV	<i>Nepovirus</i>	JF968121 (RNA2)	3 817	238	82,8	1×10^{-55}	4

5.5 VirusDetect-ohjelmalla tunnistetut virukset

VirusDetect-ohjelman tulokset vahvistivat, että poolatussa näytteessä esiintyy arabiksen mosaiikkivirusta, tupakan rattlevirusta ja lupiinin mosaiikkivirusta (taulukko 11 ja kuva 8). BLASTn-haun tulosten perusteella osa kontigeista vastasi identtisyydeltään 90,3–99-prosenttisesti ArMV:n, TRV:n ja LuMV:n referenssisekvenssejä. Linjautuneet kontigit peittivät ArMV:n, TRV:n ja LuMV:n referenssisekvenssejä 66,3–100 %:sti ja peittävyiden keskimääräinen syvyys vaihteli 100,5–315,6 välillä. Pisin ohjelmassa muodostettu kontigi oli 6120 nukleotidia pitkä, joka vastasi LuMV:ta.

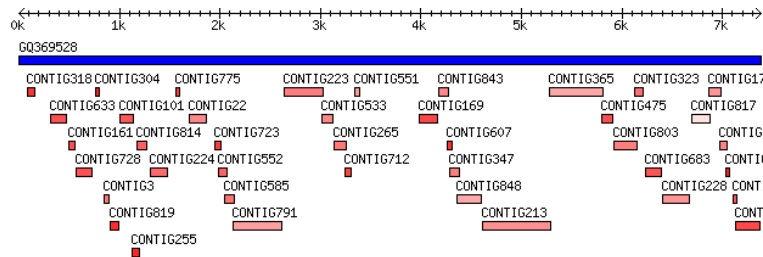
VirusDetect-ohjelman tulokset antoivat viitteitä, että näytteissä voi mahdollisesti esiintyä tuntemattomia poty-, clostero- ja hordeiviruksia (taulukko 12 ja kuva 9). BLASTn-haun tulosten perusteella osa kontigeista vastasi identtisyydeltään 83,4 %:sti pavun keltamosaiikkivirusta, referenssigenomin peittävyiden ollen 39,0 % ja peittävyiden keskimääräisen syvyyden ollen 300,9. BLASTx-haun perusteella ohran juovamosaiikkiviruksen (BSMV) RNA3-referenssisekvenssiin linjautuneet kontigit vastasivat referenssisekvenssiä 73,1-prosenttisesti, peittävyiden ollen 68,4 ja peittävyiden keskimääräisen syvyyden ollen 217,2. Minttuviruksen yhden (MV1) referenssisekvenssiin linjautuneet kontigit vastasivat referenssisekvenssiä 50-prosenttisesti, peittävyiden ollen 15,8 ja peittävyiden keskimääräisen syvyyden ollen 73,6. Juurikkaan keltavirukseen (BYV) referenssisekvenssiin linjautuneet kontigit vastasivat referenssisekvenssiä 12-prosenttisesti, peittävyiden ollen 12 ja peittävyiden keskimääräisen syvyyden ollen 86.

Taulukko 11. VirusDetect-ohjelmassa koottujen kontigien linjautuminen NCBI:n tietokannan virusten nukleotidisekvensseihin.

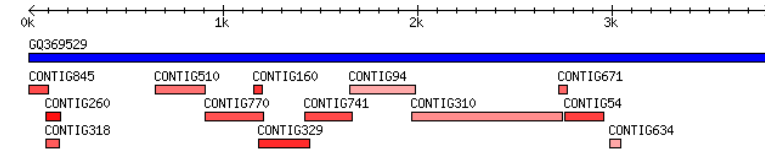
Virallinen nimi	Suomenkielinen nimi	Lyhenne	Suku	Sekvenssin yksilöllinen tunniste NCBI:ssä	Referenssi-sekvenssin pituus (nt)	Ident-tisyys (%)	Peittä-vyys (%) ¹	Peittävyuden keskimääräinen syvyys ²	Kontigien määrä	Pisin kontigi
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikkivirus	ArMv	<i>Nepovirus</i>	GQ369528 (RNA 1)	7 384	90,3	79,8	234,4	40	678
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikkivirus	ArMv	<i>Nepovirus</i>	GQ369529 (RNA2)	3 811	91,7	66,3	288,4	13	776
<i>Tobacco rattle virus</i>	Tupakan rattlevirus	TRV	<i>Tobravirus</i>	GQ903771 (RNA 1)	6 791	98,3	100	100,5	36	3 323
<i>Tobacco rattle virus</i>	Tupakan rattlevirus	TRV	<i>Tobravirus</i>	KP100069 (RNA2)	4 107	98,8	60,1	230,5	8	1 775
<i>Lupinus mosaic virus</i>	Lupiinin mosaiikkivirus	LuMV	<i>Potyvirus</i>	EU847625	10 113	99,0	99,6	315,6	8	6 120
<i>Bean yellow mosaic virus</i>	Pavun kelta-mosaiikkivirus	BYMV	<i>Potyvirus</i>	AB373203	9 535	83,4	39,7	300,9	13	444

¹ sRNA-sekvenssien kattama osuus referenssigenomista prosentteina.

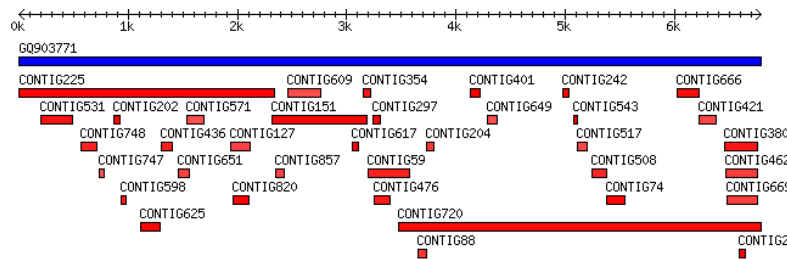
² Kuvaa montako sRNA-sekvenssin emästä vastaa tietyllä paikalla sijaitsevaa referenssigenomin emästä keskimäärin.



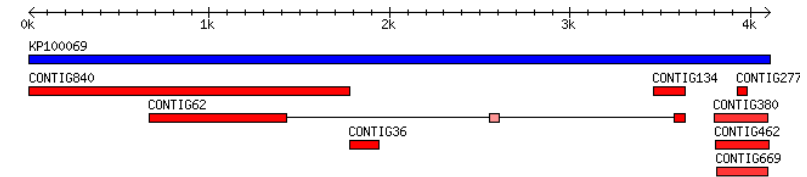
(a) Arabiksen mosaiikkivirus, RNA 1, GQ369528



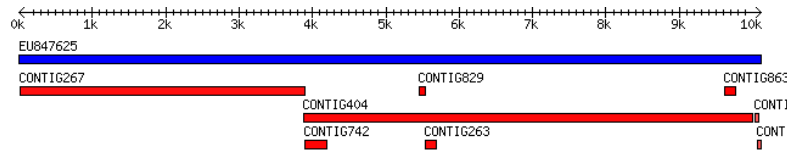
(b) Arabiksen mosaiikkivirus, RNA2, GQ36952



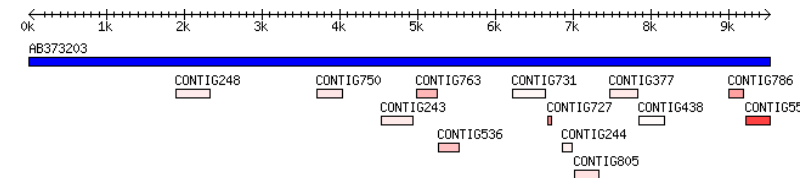
(c) Tupakan rattlevirus, RNA 1, GQ903771



(d) Tupakan rattlevirus, RNA2, KP100069



(e) Lupiinin mosaiikkivirus, EU847625



(f) Pavun keltamosaiikkivirus, AB373203

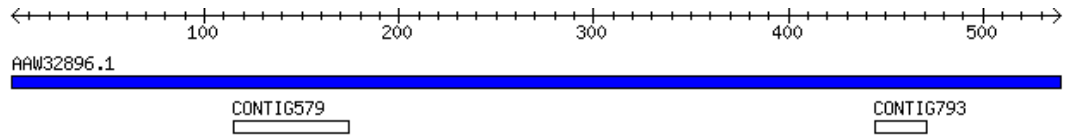
Kuva 8. Visualisointi VirusDetect-ohjelmassa koottujen kontigien linjautumisesta nukleotidisekvensseihin. Sininen linja kuvaa referenssisekvenssiä ja punaiset linjat referenssisekvenssiin linjautuneita kontigeja. Kontigin väritys kuvaa kontigin ja referenssisekvenssin välistä identtisyyttä, punaisen ollessa 100 % ja valkoisen alle 75 %. Kuvien alla olevat koodit viittaavat referenssisekvenssin tunnisteeseen NCBI:n tietokannassa.

Taulukko 12. VirusDetect-ohjelmassa koottujen kontigien linjautuminen NCBI:n tietokannan virusten proteiinisekvensseihin.

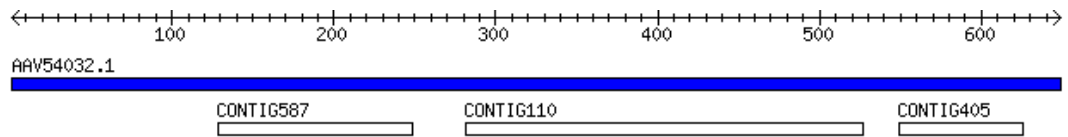
Virallinen nimi	Suomenkielinen nimi	Lyhenne	Suku	Sekvenssin yksilöllinen tunniste NCBI:ssä	Referenssisekvenssin pituus aminohappoina	Identisyys (%)	Peittävyys (%) ¹	Peittävyuden keskimääräinen syvyys ²	Kontigien määrä	Pisin kontigi
<i>Barley stripe mosaic virus</i>	Ohran juovamosaiikkivirus	BSMV	<i>Hordeivirus</i>	AAV54032.1 (RNA3)	648	73,1	68,4	217,2	3	246
<i>Bean yellow mosaic virus</i>	Pavun keltamosaiikkivirus	BYMV	<i>Potyvirus</i>	BAF96025.1	3 056	77,2	31,6	425,4	7	407
<i>Beet yellows virus</i>	Juurikkaan keltavirus	BYV	<i>Closterovirus</i>	CAA51871.1	3 074	61,1	12	86,0	10	59
<i>Mint virus 1</i>	Minttuvirus 1	MV1	<i>Closterovirus</i>	AAW32896.1	15 450	50,0	15,8	73,6	2	64

¹ sRNA-sekvenssien kattama osuus referenssigenomista prosentteina.

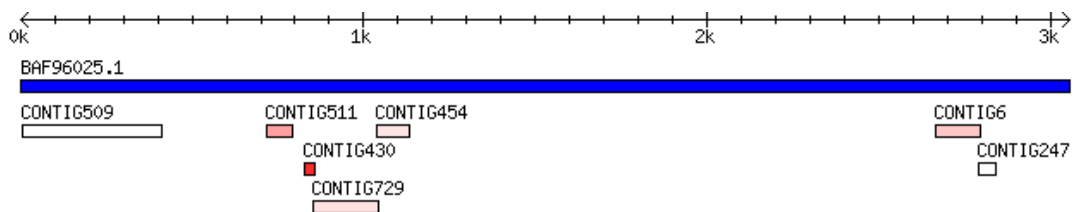
² Kuvaa montako sRNA-sekvenssin emästä vastaa tietyllä paikalla sijaitsevaa referenssigenomin emästä keskimäärin.



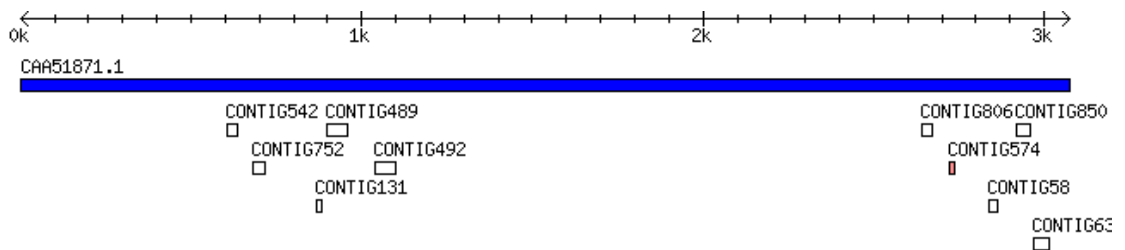
(a) Minttuvirus 1, NCBI:n referenssitietokannan tunniste AAW32896



(b) Ohran juovamosaiikkivirus, RNA3, NCBI:n referenssitietokannan tunniste AAV54032



(c) Pavun keltamosaiikkivirus, NCBI:n referenssitietokannan tunniste BAF96025



(d) Juurikkaan keltavirus, NCBI:n referenssitietokannan tunniste CAA51871

Kuva 9. Visualisointi VirusDetect-ohjelmassa koottujen kontigien linjautumisesta proteiinisekvensseihin. Sininen linja kuvaa referenssisekvenssiä ja punaiset ja valkoiset linjat referenssisekvenssiin linjautuneita kontigeja. Kontigin väritys kuvaa kontigin ja referenssisekvenssin välistä identtisyyttä, punaisen ollessa 100 % ja valkoisen alle 75 %.

5.6 Bowtie 2 -ohjelmalla referenssiviruksiin linjatut sRNA-sekvenssit

Bowtie-tulosten perusteella peittävydessä ja peittävyyden keskimääräisessä syvyydessä esiintyi referenssisekvenssien välillä vaihtelua (taulukko 13). ArMV:n, TRV:n ja LuMV:n referenssisekvenssien kohdalla peittävyys oli useita kymmeniä prosentteja (59,2 %–99,9 %) ja peittävyyden keskimääräinen syvyys useita kymmeniä tai satoja nukleotideja (66,5–308,9). BYMV:n, SBMV:n, CIYVV:n ja GFLV:n referenssisekvenssien peittävyys vaihteli 2,8–18,1 %:n välillä ja peittävyyden keskimääräinen syvyys välillä 2,3–16,2. Ohjelman tulokset eivät anna viitteitä hordei- tai closterovirusten esiintymisestä poolatussa näytteessä, sillä kyseisiin sukuihin kuuluvien referenssisekvenssien peittävyys ja peittävyyden syvyys jäivät alhaiseksi (peittävyys 0,14–3,3 % ja peittävyyden keskimääräinen syvyys 0,003–1,6).

ArMV:n RNA1-referenssisekvenssin peittävyyden syvyys oli suurin RNA-helikaaasia koodaavalla alueella (kuva 14) ja ArMV:n RNA2-referenssisekvenssin kohdalla liikkumisproteiinia koodaavalla alueella (kuva 15). sRNA-sekvenssit peittivät TRV:n RNA1-referenssisekvenssiä lähes sataprosenttisesti, mutta TRV:n RNA2-referenssisekvenssistä kokonaan ilman peittävyyttä jäivät osat 2c- ja 2b-proteiinia koodaavista alueista (kuvat 12 ja 13).

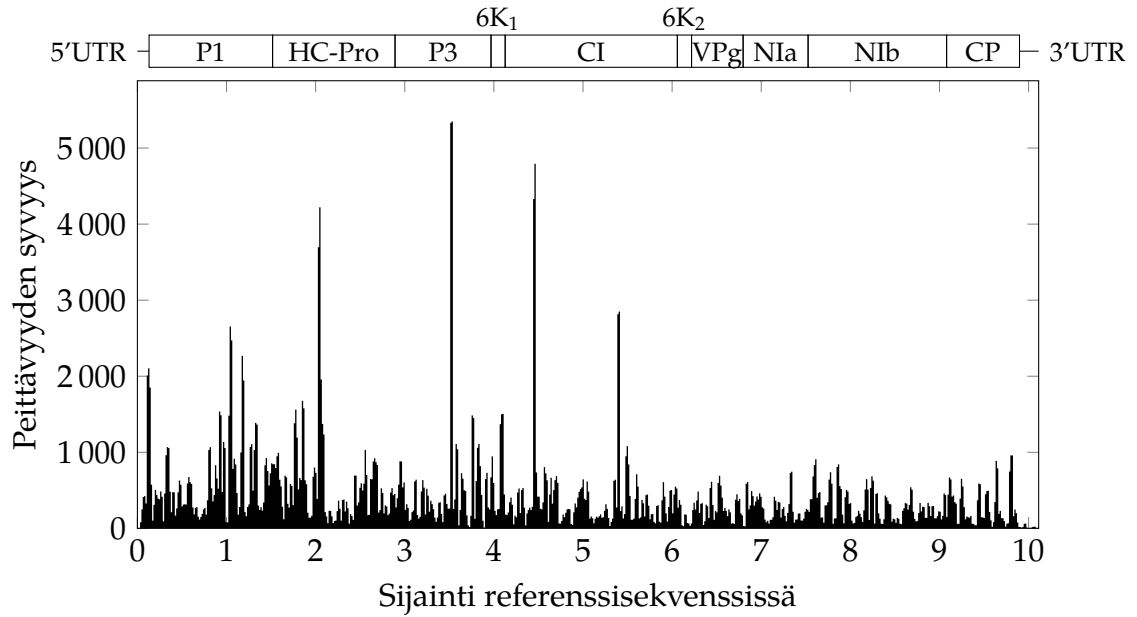
sRNA-sekvenssit peittivät LuMV:n referenssisekvenssiä lähes sataprosenttisesti. Peittävyyden syvyys oli suurempi LuMV-referenssigenomin 5'-päässä kuin 3'-päässä (kuva 10). BYMV:n kohdalla peittävyys ja peittävyyden syvyys osuivat eniten niille genomin alueille, jotka koodaavat CP-, NIb-, CI- tai HCpro-proteiinia (kuva 11). SMV:n ja CIYVV:n kohdalla peittävyys ja peittävyyden syvimmät kohdat osuivat lähinnä niille genomin alueille, jotka koodaavat CP- tai NIb-proteiinia ja CIYVV:n tapauksessa myös CI-proteiinia (kuva 16). GFLV:n referenssisekvenssien peittävyys ja peittävyyden syvyys olivat parhaimpia RNA2:n 5'-pään alussa, jossa esiintyvät UTR-alue ja P2-proteiinia koodaava alue (kuva 16).

Taulukko 13. sRNA-sekvenssien linjautuminen referenssisekvensseihin. Tulokset ovat Bowtie-ajosta, jossa ei sallittu yhtään virhettä linjauksessa (parametri N = 0).

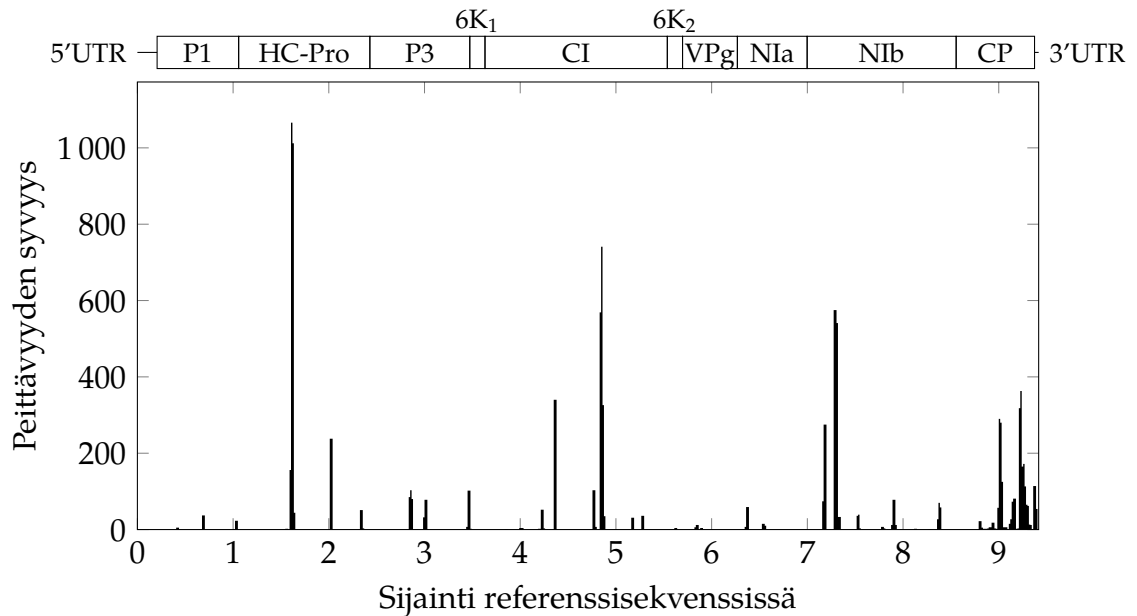
Virallinen nimi	Suomenkielinen nimi	Lyhenne	Suku	Sekvenssin yksilöllinen tunnistus NCBI:ssä	Referenssigenomin pituus (nt)	Peittävyys (%) ¹	Peittävyys keskimääräinen syvyys ²
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikkivirus (RNA1)	ArMV	<i>Nepovirus</i>	EU617326	7 334	59,2	66,5
<i>Arabidopsis mosaic virus</i>	Arabiksen mosaiikkivirus (RNA2)	ArMV	<i>Nepovirus</i>	EU617327	3 812	59,5	80,1
<i>Tobacco etch virus</i>	Tupakan rattlevirus (RNA1)	TRV	<i>Tobravirus</i>	GQ903771	6 791	99,9	105,2
<i>Tobacco etch virus</i>	Tupakan rattlevirus (RNA2)	TRV	<i>Tobravirus</i>	KJ826366	4 154	60,8	120,1
<i>Lupinus mosaic virus</i>	Lupiinin mosaiikkivirus	LuMV	<i>Potyvirus</i>	EU847625	10 113	98,7	308,9
<i>Bean yellow mosaic virus</i>	Pavun keltamosaiikkivirus	BYMV	<i>Potyvirus</i>	HG970867	9 417	15,1	12,7
<i>Soybean mosaic virus</i>	Soijapavun mosaiikkivirus	SMV	<i>Potyvirus</i>	KP710868	9 994	2,8	2,3
<i>Grapevine fanleaf virus</i>	Viiniköynnöksen viuhkalehtivirus (RNA1)	GFLV	<i>Nepovirus</i>	JF968120	7 342	9,4	7,5
<i>Grapevine fanleaf virus</i>	Viiniköynnöksen viuhkalehtivirus (RNA2)	GFLV	<i>Nepovirus</i>	JF968121	3 817	18,1	16,2
<i>Clover yellow vein virus</i>	Apilan keltasuonivirus	CIYVV	<i>Potyvirus</i>	HG970870	9 439	8,1	6,2
<i>Beet yellows virus</i>	Juurikkaan keltavirus	BYV	<i>Closterovirus</i>	NC_001598	15 480	0,14	0,003
<i>Mint virus 1</i>	Minttuivirus 1	MV1	<i>Closterovirus</i>	NC_006944	15 450	0,14	0,016
<i>Barley stripe mosaic virus</i>	Ohran juovamosaiikkivirus (RNA1)	BSMV	<i>Hordeivirus</i>	NC_003469	3 768	1,3	0,1
<i>Barley stripe mosaic virus</i>	Ohran juovamosaiikkivirus (RNA2)	BSMV	<i>Hordeivirus</i>	NC_003481	3 289	0,7	0,2
<i>Barley stripe mosaic virus</i>	Ohran juovamosaiikkivirus (RNA3)	BSMV	<i>Hordeivirus</i>	NC_003478	3 164	3,3	1,6

¹ sRNA-sekvenssien kattama osuus referenssigenomista prosentteina.

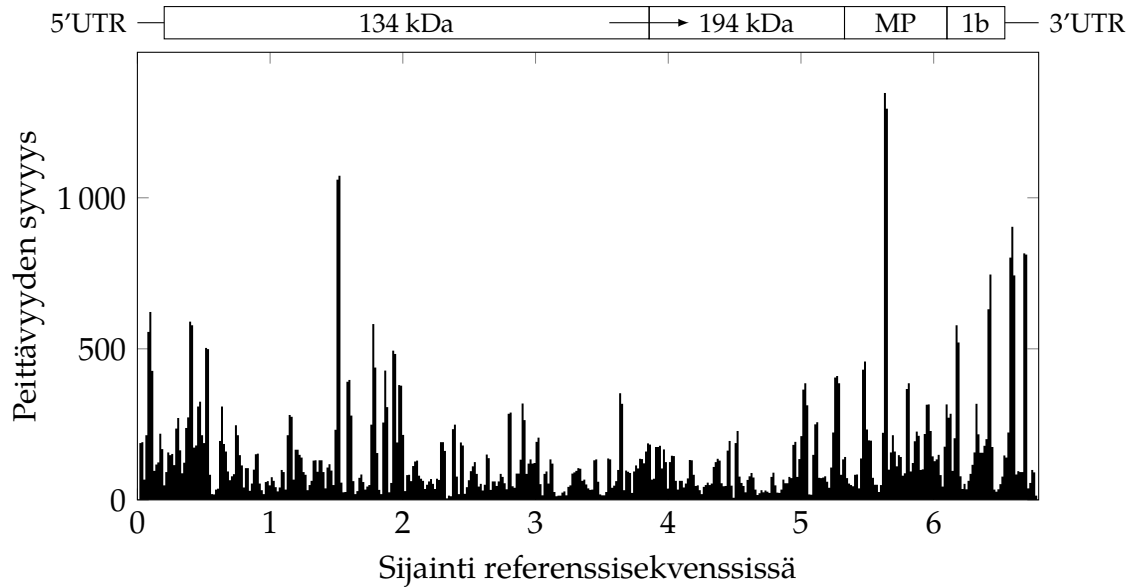
² Kuvaa montako sRNA-sekvenssin emästä vastaa tietyllä paikalla sijaitsevaa referenssigenomin emästä keskimäärin.



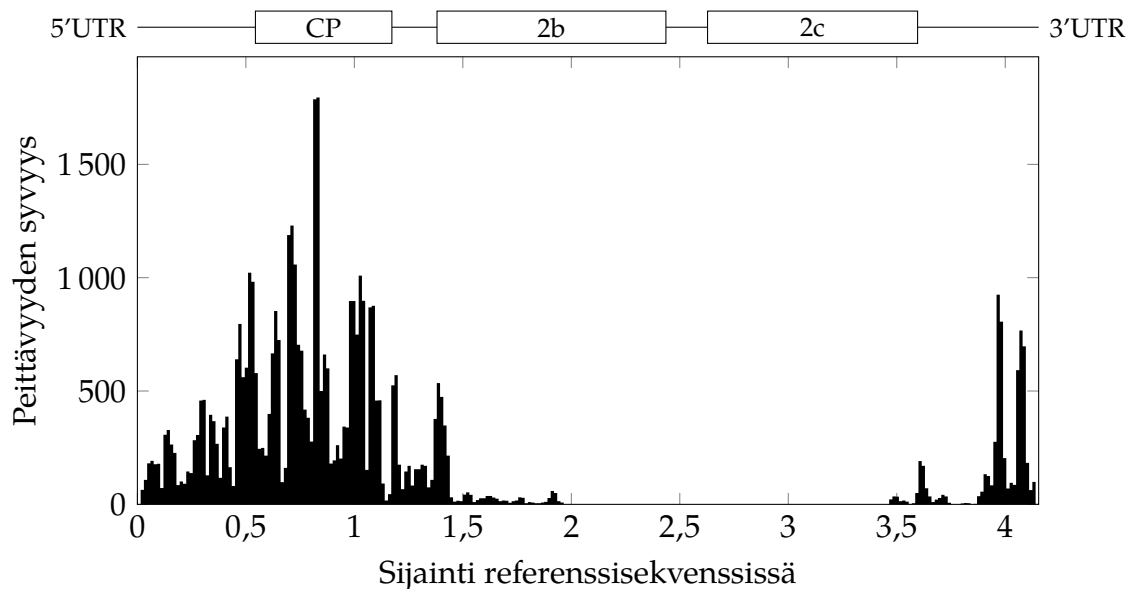
Kuva 10. sRNA-sekvenssien linjautuminen lupiinin mosaiikkivirukseen (LuMV) (NCBI:n referenssitietokannan tunniste EU847625). Y-akseli kuvaa peittävyiden syvyyttä nukleotideina ja x-akseli peittävyiden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella.



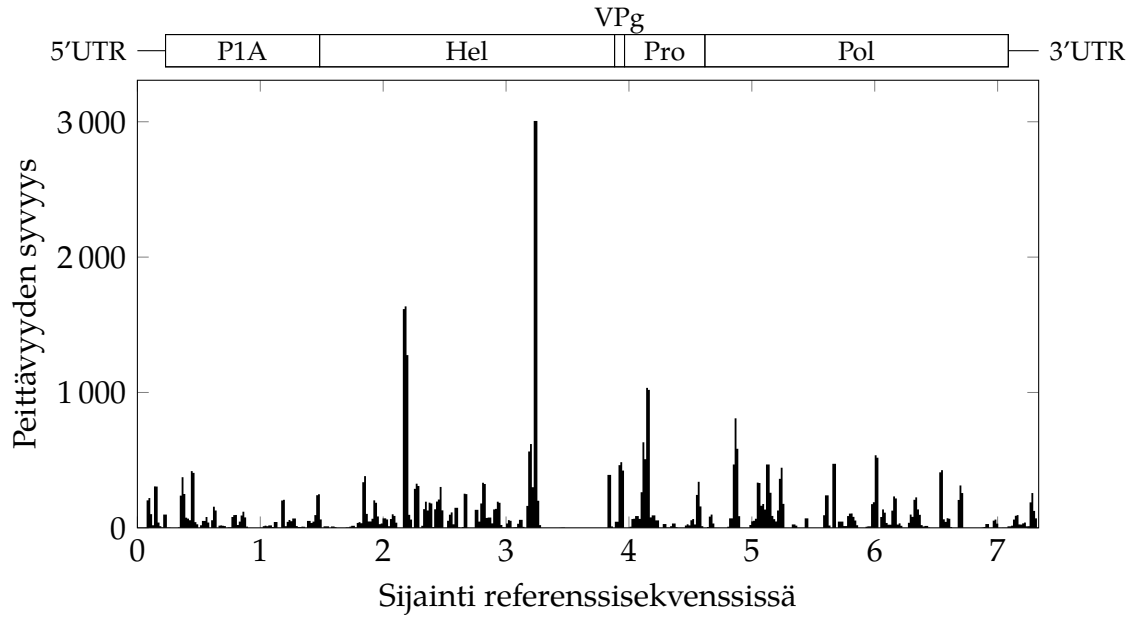
Kuva 11. sRNA-sekvenssien linjautuminen pavun keltamosaiikkivirukseen (BYMV) (NCBI:n referenssitietokannan tunniste HG970867). Y-akseli kuvaa peittävyiden syvyyttä nukleotideina ja x-akseli peittävyiden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella.



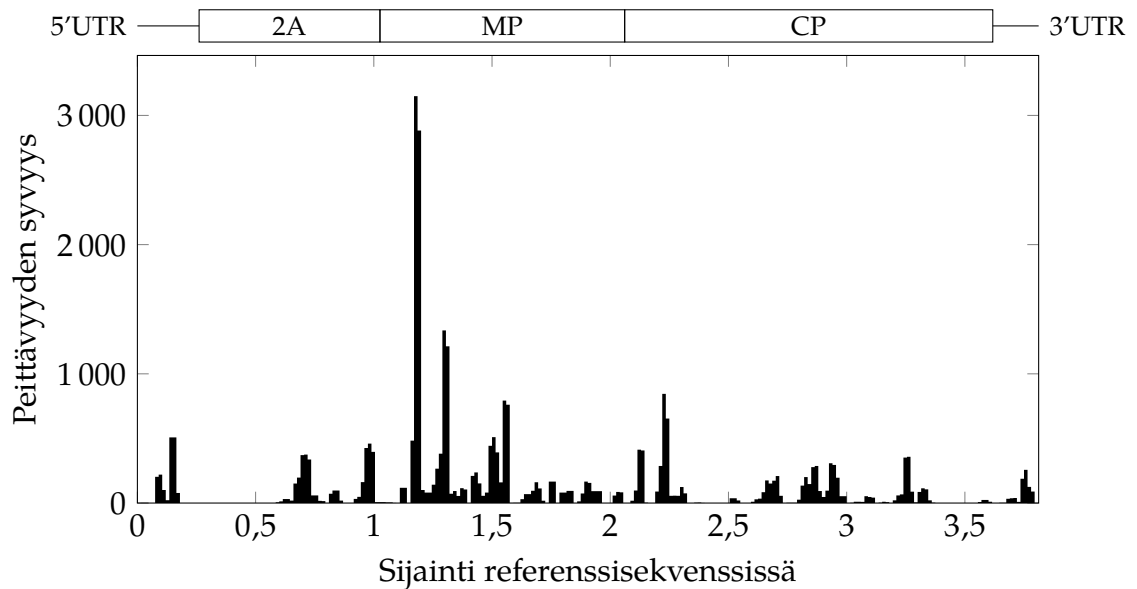
Kuva 12. sRNA-sekvenssien linjautuminen tupakan rattleviruksen (TRV) RNA1-referenssisekvenssiin (NCBI:n referenssitietokannan tunniste GQ903771). Y-akseli kuvaa peittävyden syvyyttä nukleotideina ja x-akseli peittävyden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella. Kuvassa näkyvä nuoli tarkoittaa, että proteiinisynteesi jatkuu lopetuskodonin jälkeen.



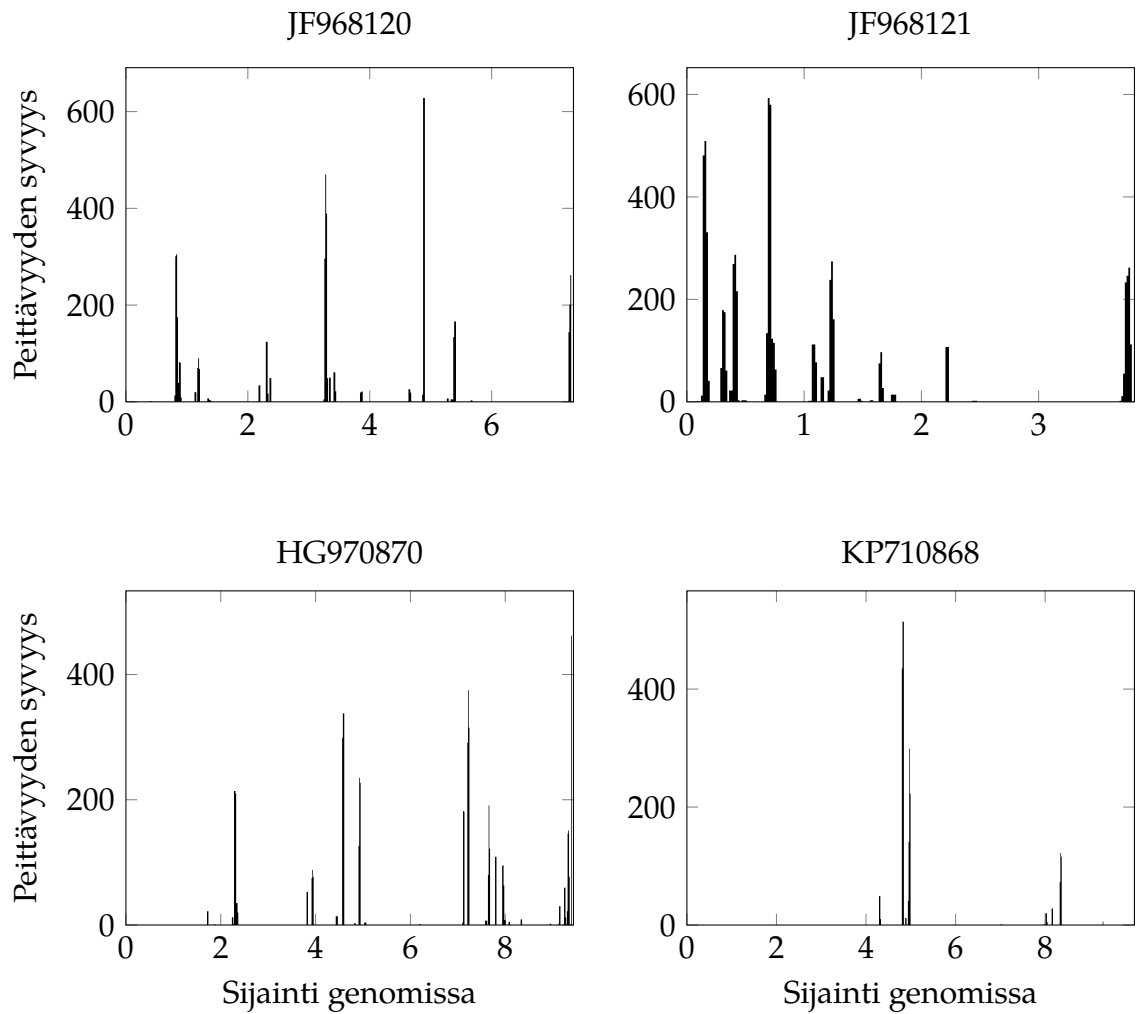
Kuva 13. sRNA-sekvenssien linjautuminen tupakan rattleviruksen (TRV) RNA2-referenssisekvenssiin (NCBI:n referenssitietokannan tunniste KJ826366). Y-akseli kuvaa peittävyden syvyyttä nukleotideina ja x-akseli peittävyden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella.



Kuva 14. sRNA-sekvenssien linjautuminen arabiksen mosaiikkiviruksen (ArMV) RNA1-referenssisekvenssiin (NCBI:n referenssitietokannan tunniste EU617326). Y-akseli kuvaa peittävyden syvyyttä nukleotideina ja x-akseli peittävyden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella.



Kuva 15. sRNA-sekvenssien linjautuminen arabiksen mosaiikkiviruksen (ArMV) RNA2-referenssisekvenssiin (NCBI:n referenssitietokannan tunniste EU617327). Y-akseli kuvaa peittävyden syvyyttä nukleotideina ja x-akseli peittävyden syvyyden sijaintia nukleotideina referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia. Referenssigenomin rakenne näkyy kuvan yläpuolella.



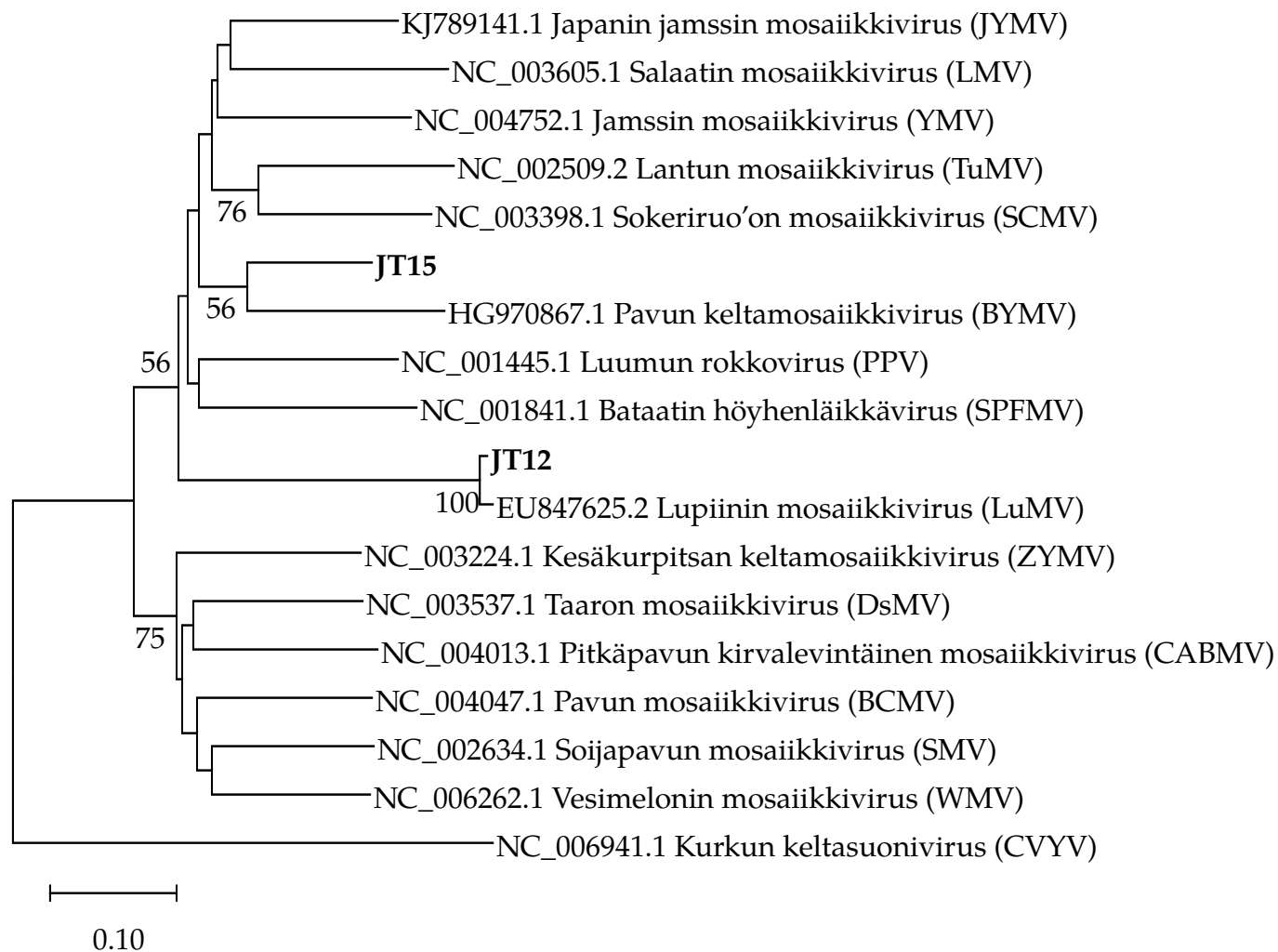
Kuva 16. sRNA-sekvenssien linjautuminen viiniköynnöksen viuhkalehtiviruksen (GFLV) RNA1 (NCBI:n referenssitietokannan tunniste JF968120) ja viiniköynnöksen viuhkalehtiviruksen RNA2 (NCBI:n referenssitietokannan tunniste JF968121), apilan keltasuoniviruksen (CIYVV) (NCBI:n referenssitietokannan tunniste HG970870) ja soijapavun mosaiikkiviruksen (SMV) (NCBI:n referenssitietokannan tunniste KP710868) referenssisekvensseihin. Y-akseli kuvaa peittävyyden syvyyttä ja x-akseli peittävyyden syvyyden sijaintia referenssigenomissa. X-akselin luvut vastaavat tuhatta emäsparia.

5.7 Sekvenssien fylogeneettinen analyysi

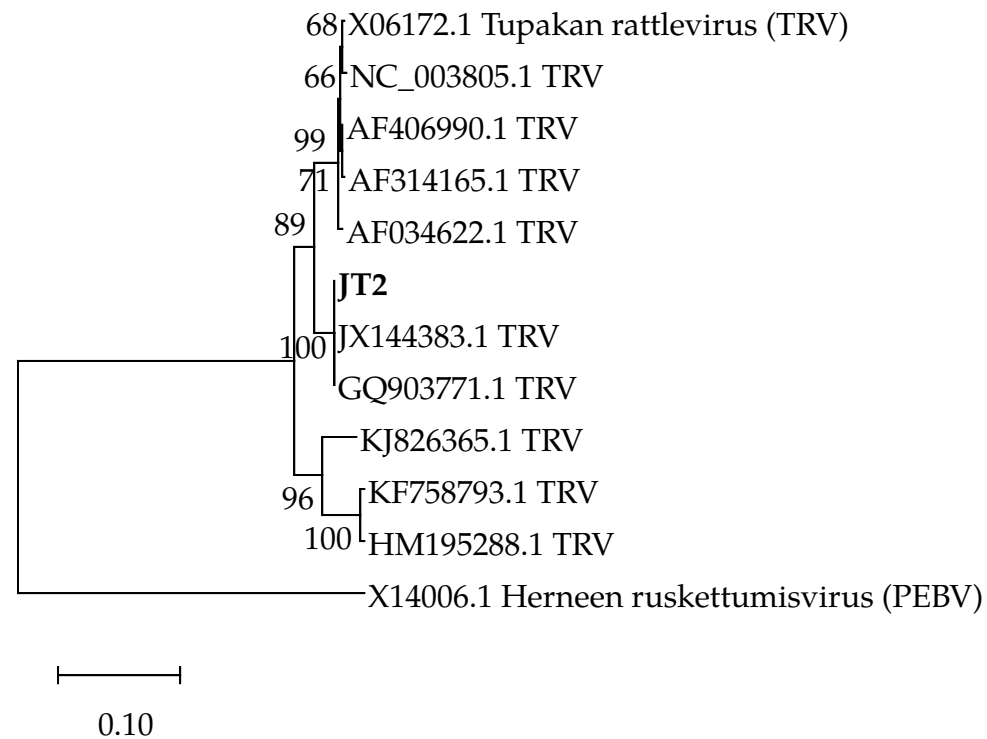
Fylogeneettisessä analyysissä tarkasteltiin näytteiden JT12 ja JT15 sukulaisuussuhteita eri potyviruslajeihin NIb-proteiinia koodaavan osittaisen alueen perusteella (liite C, taulukko 1). Fylogeneettisen puun rakentamisessa käytetyn sekvenssialueen pituus oli 314 nukleotidia. Näyte JT12 (lupiini) muodosti fylogeneettisessä puussa sukuhaaran LuMV:n kanssa ja näyte JT15 (valkomesikkä) BYMV:n kanssa (kuva 17). Eri potyviruslajeista näyte JT12 vastasi nukleotidisekvenssiltään eniten LuMV:ta (identtisyys 98,4 %) ja näyte JT15 eniten japanin jamssin mosaiikkivirusta (JYMV) (identtisyys 66,4 %) (taulukko 14).

Näytteen JT2 sukulaisuussuhdetta tarkasteltiin eri TRV-isolaatteihin RdRp:tä koodaavan osittaisen alueen perusteella (liite C, taulukko 2). Fylogeneettisen puun rakentamisessa käytetyn sekvenssialueen pituus oli 721 nukleotidia. Näyte JT2 vastasi nukleotidisekvenssiltään eniten TRV-isolaatteja X144383.1 ja GQ903771.1 (identtisyys 99,6 %) (taulukko 15), joiden kanssa näyte muodosti fylogeneettisessä puussa oman sukuhaaran (kuva 18).

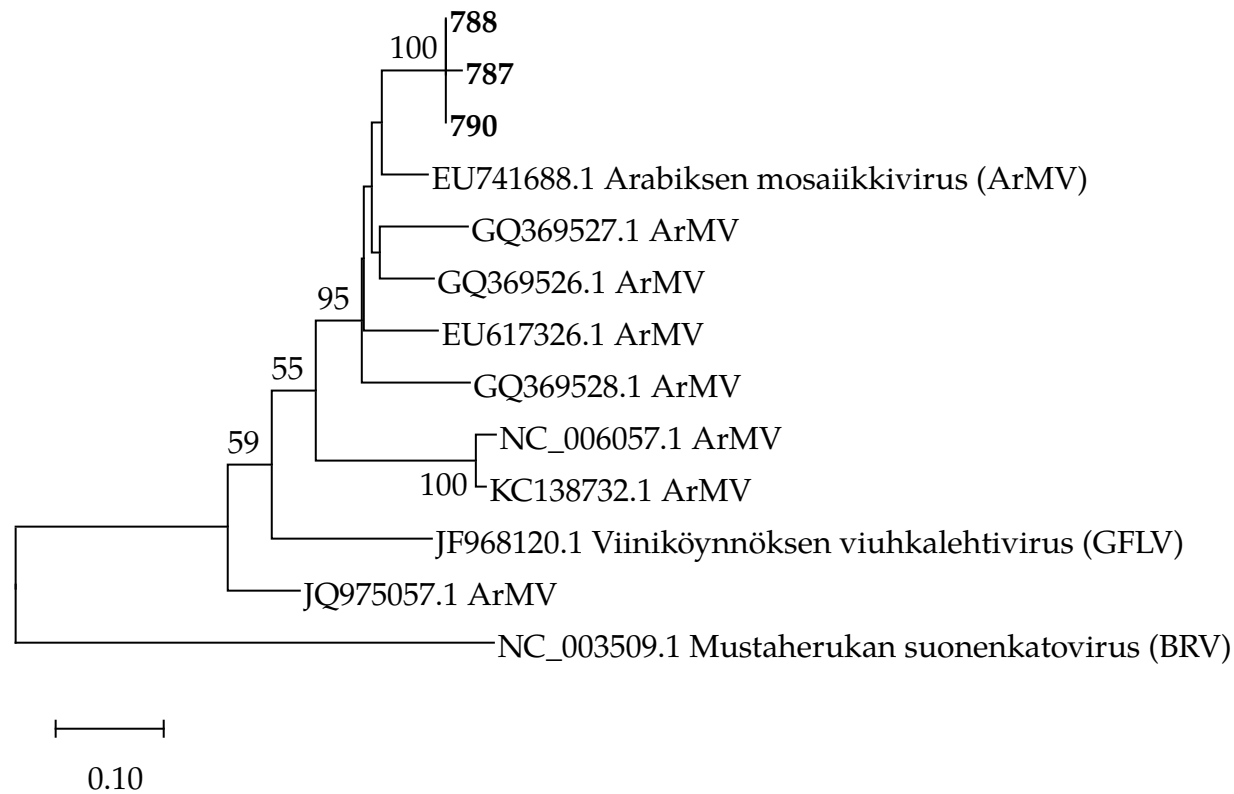
Näytteiden 787, 788 ja 790 sukulaisuussuhteita tarkasteltiin eri ArMV-isolaatteihin ja yhteen GFLV-isolaattiin RdRp:tä koodaavan osittaisen alueen perusteella (liite C, taulukko 3). Fylogeneettisen puun rakentamisessa käytetyn sekvenssialueen pituus oli 315 nukleotidia. Fylogeneettisessä puussa näytteet muodostivat oman sukuhaaran (kuva 19). Näytteiden kanssa lähimmän sukulaissuhteen muodosti puussa Uusi-Seelannista löydetty ja kvinoasta eristetty isolaatti EU741688.1. Kyseinen isolaatti vastasi referenssisekvensseistä eniten näytteitä, identtisyyden ollen 88,6–90,7 % (taulukko 16).



Kuva 17. Näytteiden JT12 ja JT15 sekä eri potyvirusreferenssisekvenssien fylogeneettinen ryhmittely NlB-proteiinia koodaavan osittaisen nukleotidialueen (314 nt) perusteella. Puussa esitetään bootstrap-arvot, jotka ovat saaneet arvoksi yli 50. Mittakaava osoittaa kymmentä nukleotidimuutosta sadasta nukleotidista. Ulkoryhmänä toimii ipomovirusiin luokiteltava kurkun keltasuonivirus.



Kuva 18. Näytteen JT2 ja eri TRV-isolaattien fylogeneettinen ryhmittely RdRp:tä koodaavan osittaisen nukleotidialueen (721 nt) perusteella. Puussa esitetään bootstrap-arvot, jotka ovat saaneet arvoksi yli 50. Mittakaava osoittaa kymmentä nukleotidimuutosta sadasta nukleotidista. Ulkoryhmänä toimii tobraviruksiin luokiteltava herneen ruskettumisvirus.



Kuva 19. Näytteiden 787, 788 ja 790, GFLV-isolaatin ja eri ArMV-isolaattien fylogeneettinen ryhmittely RdRp:tä koodaavan osittaisen nukleotidialueen (315 nt) perusteella. Puussa esitetään bootstrap-arvot, jotka ovat saaneet arvoksi yli 50. Mittakaava osoittaa kymmentä nukleotidimuutosta sadasta nukleotidista. Ulkoryhmänä toimii nepoviruksen alaryhmä c:hen luokiteltava mustaherukan suonenkatovirus.

Taulukko 14. Potyvirusreferenssisekvenssien identtisyys näytteisiin JT12 ja JT15 NIb-proteiinia koodaavan osittaisen nukleotidialueen (314 nt) perusteella.

Nimi	Lyhenne	Sekvenssin yksilöllinen tunniste NCBI:ssä	Identtisyys-% JT12:een verrattuna	Identtisyys-% JT15:een verrattuna
JT12			100,0	56,9
JT15			56,9	100,0
Bataatin höyhenläikkävirus	SPFMV	NC_001841.1	67,0	58,1
Jamssin mosaiikkivirus	YMV	NC_004752.1	65,8	60,2
Japanin jamssin mosaiikkivirus	JYMV	KJ789141.1	63,5	66,4
Kesäkurpitsan keltamosaiikkivirus	ZYMV	NC_003224.1	65,1	56,6
Lantun mosaiikkivirus	TuMV	NC_002509.2	66,1	57,7
Lupiinin mosaiikkivirus	LuMV	EU847625.2	98,4	57,3
Luumun rokkovirus	PPV	NC_001445.1	68,0	59,6
Pavun keltamosaiikkivirus	BYMV	HG970867.1	63,2	62,8
Pavun mosaiikkivirus	BCMV	NC_004047.1	59,2	55,4
Pitkäpavun kirvalevintäinen mosaiikkivirus	CABMV	NC_004013.1	65,4	58,4
Salaatin mosaiikkivirus	LMV	NC_003605.1	61,6	60,6
Soijapavun mosaiikkivirus	SMV	NC_002634.1	65,1	56,6
Sokeriruo'on mosaiikkivirus	SCMV	NC_003398.1	65,8	58,8
Taaron mosaiikkivirus	DsMV	NC_003537.1	65,4	58,8
Vesimelonin mosaiikkivirus	WMV	NC_006262.1	64,1	55,4

Taulukko 15. Tupakan rattlevirus -isolaattien (TRV) identtisyys näytteeseen JT2 RdRp:tä koodaavan osittaisen nukleotidialueen (721 nt) perusteella.

Sekvenssin yksilöllinen tunniste NCBI:ssä	Identtisyys-% JT2:een verrattuna
GQ903771.1	99,6
JX144383.1	99,6
AF406990.1	95,8
AF034622.1	95,7
AF314165.1	95,7
X06172.1	95,7
NC_003805.1	95,3
KJ826365.1	92,4
KF758793.1	91,5
HM195288.1	91,4

Taulukko 16. Arabiksen mosaiikkivirus -isolaattien (ArMV) ja yhden viiniköynnöksen viuhkalehtivirus -isolaatin (GFLV) identtisyys näytteisiin 787, 788 ja 790 RdRp:tä koodaavan osittaisen nukleotidialueen (315 nt) perusteella.

Koodi-numero tai lyhenne	Sekvenssin yksilöllinen tunniste NCBI:ssä	Identtisyys-% 787:ään verrattuna	Identtisyys-% 788:aan verrattuna	Identtisyys-% 790:een verrattuna
787		100,0	96,6	96,6
788		96,6	100,0	100,0
790		96,6	100,0	100,0
ArMV	EU741688.1	88,6	90,7	90,7
ArMV	GQ369526.1	87,9	88,7	88,7
ArMV	EU617326.1	85,9	88,4	88,4
ArMV	GQ369527.1	84,2	86,1	86,1
ArMV	GQ369528.1	82,6	85,1	85,1
GFLV	JF968120.1	75,2	76,7	76,7
ArMV	JQ975057.1	74,8	76,7	76,7
ArMV	KC138732.1	74,5	75,4	75,4
ArMV	NC_006057.1	74,8	75,4	75,4

5.8 Yhteenveto multiplex-RT-PCR:llä, Velvet- ja VirusDetect-ohjelmilla tunnistetuista viruksista

Multiplex-RT-PCR sekä siRNA-syväseksensoinnissa käytetyt Velvet- ja VirusDetect-ohjelmat vahvistivat kokeen kasviaineistossa esiintyvän LuMV:ta, TRV:ta ja ArMV:ta (taulukko 17). Kaikki käytetyt menetelmät viittasivat kasviaineistossa esiintyvän potyvirusukuun kuuluvaa lajia, jolle löydettiin tietokannasta vain osittaisia vastaavuuksia. Muista menetelmistä poiketen VirusDetectin tulokset antoivat osittaisia vastaavuuksia hordei- ja closteroviruksiin. Velvet antoi muista menetelmistä poiketen osittaisia vastaavuuksia GFLV:een, mutta koko genomien sijaan ohjelman tulokset löysivät vastaavuuksia vain kontigien ja GFLV:n RNA2-referenssisekvenssin välille.

Taulukko 17. Vertailu multiplex-RT-PCR:llä, Velvet- ja VirusDetect-ohjelmalla havaituista viruksista. Merkintä + ilmaisee mikäli menetelmä havaitsi viruksen.

Lajinimi	Lyhenne	Suku	Multiplex-RT-PCR	Velvet	VirusDetect
Lupiinin mosaiikkivirus	LuMV	<i>Potyvirus</i>	+	+	+
Arabiksen mosaiikkivirus	ArMV	<i>Nepovirus</i>	+	+	+
Tupakan rattlevirus	TRV	<i>Tobravirus</i>	+	+	+
Tuntematon potyvirus	-	<i>Potyvirus</i>	+	+	+
Tuntematon nepovirus	-	<i>Nepovirus</i>		+	
Tuntematon clostero-virus	-	<i>Closterovirus</i>			+
Tuntematon hordeivirus	-	<i>Hordeivirus</i>			+

6 Tulosten tarkastelu

Tässä tutkimuksessa haluttiin luoda kaksi tai useampi multiplex-RT-PCR-testi, joilla voidaan havaita yhtäaikaaisesti kaikkia niitä RNA-viruksia, joita kokemuksen ja kirjallisuuden perusteella on todennäköisintä löytää Eviraan (nyk. Ruokavirasto) tarkastettavaksi tulevasta kasviaineistosta. Alukeparit valittiin kirjallisuuden perusteella, koska siten saatiin käyttöön tiettyä virusheimoa, -sukua tai -alaryhmää tunnistavat alukeparit, joiden herkkyys ja valikoivuus on valmiiksi testattu ja jotka todistetusti tunnistavat suuren joukon tiettyyn virusryhmään kuuluvia viruksia. Samaa lähestymistapaa multiplex-PCR-testien kehittämiseen on käytetty onnistuneesti esimerkiksi Gambinon ja Gribaudon (2006) sekä Wein ym. (2009) tekemissä tutkimuksissa.

Tutkimuksessa luotiin kaksi multiplex-RT-PCR-testiä, jotka tunnistavat yhteensä kahdeksaan eri virusryhmään kuuluvia viruksia. Optimoitujen multiplex-RT-PCR-testien toimivuus testattiin sekä positiivisilla kontrolliviruksilla että luonnollisesti infektoiduneilla kasvinäytteillä. Testien herkkyys testattiin positiivisista viruskontrolleista tehdyillä laimennoksilla. Kehitettyjä multiplex-RT-PCR-testejä voidaan pitää poikkeuksellisina, sillä monessakaan laboratoriossa ei ole käytettävissä yhtä laaja-alaista testiä RNA-virusten tunnistamiseen kuin mitä tutkimuksessa optimoiduilla testeillä voidaan tunnistaa. Multiplex-PCR säästää aikaa sekä materiaalikustannuksia verrattuna simplex-PCR:ään, mistä johtuen multiplex-PCR:n käyttö tarjoaa selviä etuja Eviran kasvintuhoojalaboratoriolle, jossa käsitellään suuria näytemääriä nopeassa aikataulussa.

Tutkimustulokset osoittivat multiplex-RT-PCR-testien kykenevän tunnistamaan kaikki positiiviset viruskontrollit kun infektoiduneen kasvin kokonais-RNA:n konsentraatio oli alle 5 ng/μl, mitä Eviran kasvintuhoojalaboratoriossa pidetään riittävänä multiplex-PCR-testin herkkyyden kannalta. Geelielektroforeesikuvista on havaittavissa oikean kokoisten tuotteiden lisäksi epäspesifisiä tuotteita sekä kohdesekvenssien epätasaista monistumista, mistä johtuen multiplex-testien lisäoptimointi kuuluu jatkotutkimuksen aiheisiin. PCR-testien suorittamisessa käytettyjen reagenssien valmistaja suosittelee kokeilemaan eri magnesiumpitoisuuksia ja alukkeiden sitoutumislämpötiloja, mikäli epäspesifisiä tuotteita on havaittavissa. Epäspesifiset tuotteet kuluttavat reaktioseoksen reagensseja jättäen vähemmän reagensseja kohdesekvenssien monistumiseen (Henson ja French 1993), mikä voi johtaa testissä väärin negatiivisiin tuloksiin. Optimoidut multiplex-PCR-testit monistivat oikean kokoiset tuotteet positiivisista kontrolliviruksista jokaisella testikerralla, mikä lisää luottamusta testien toistettavuutta kohtaan.

Teoriassa yksi mahdollinen ratkaisu kohdesekvenssien epätasaiseen monistumiseen on alentaa tehokkaasti monistavien ja kasvattaa heikosti monistavien alukeparien pitoisuutta. Tosin Wei ym. (2009) muistuttavat, että kasviviruset voivat esiintyä kasveissa vaihtelevina pitoisuuksina. Tästä johtuen heidän mukaan erilaisien konsentraatioiden käyttöä eri alukepareille samassa reaktiossa ei yleensä voida pitää varteenotettavana ratkaisuna vähentämään kohdesekvenssien epätasaista monistumista.

Geelikuvat multiplex-RT-PCR:ssä testatuista kasvinäytteistä (kuvat 6 ja 7) osoittavat epäspesifisten tuotteiden määrän vaihdelleen riippuen siitä mikä testattu kasvilaji oli kyseessä. Tulosten perusteella vaikuttaa mahdolliselta, että osa kasvilajeista tuottaa multiplex-RT-PCR-testissä epäspesifisiä tuotteita johtuen alukkeiden sitoutumisesta virussekvenssin sijaan isäntäkasvin sekvenssiin. Tulosten perusteella multiplex-RT-PCR-testien soveltuvuus virusten testaamiseen tietystä kasvilajista on selvitettävä ennen varsinaista rutiinianalyysiä.

Saadut tutkimustulokset toimivat lähtökohtana multiplex-RT-PCR-testien validoinnille, jonka avulla voidaan varmistua menetelmien soveltuvuudesta käyttö-tarkoitukseensa ja menetelmien antamien tulosten luotettavuudesta. Tässä tutkimuksessa ei testattu multiplex-RT-PCR-testien uusittavuutta, mikä on tehtävä testien validoinnissa. Tutkimuksen tulosten perusteella siRNA-syväsekvenssoinnin ja multiplex-PCR:n kyky tunnistaa viruksia oli samanlainen, mikä vahvistaa luotamusta optimoituja multiplex-PCR-testejä kohtaan. Tämän tutkimuksen jälkeen optimoidut multiplex-PCR-testit ovat validoitu, eli akkreditoitu, ja validointi on tarkastettu ulkopuolisella asiantuntijalla, mikä varmistaa menetelmän soveltuvan rutiinidiagnostiikan käyttöön (Johanna Santala, Ruokavirasto, sähköpostiviesti kirjoittajalle 19.12.2018).

Tulosten perusteella siRNA-syväsekvenssointi tunnisti osasta löytyneistä viruksista lähes niiden koko genomin osittain hajonneesta RNA:sta, mikä puoltaa menetelmän soveltuvan virusten laaja-alaiseen testaamiseen. Nykyään siRNA-syväsekvenssointia käytetään joissakin laboratorioissa osana rutiinidiagnostiikkaa, mutta on huomioitava ettei menetelmää ole vielä akkreditoitu yhdessäkään kasvilaboratoriossa. Toistaiseksi yksikään kasvintarkastusviranomainen ei hyväksy sitä vielä ainoana menetelmänä kasvivirusien tunnistamiseen, vaan siRNA-syväsekvenssoinnista saatavat tulokset ovat aina vahvistettava vielä muulla tunnistusmenetelmällä (Johanna Santala, Ruokavirasto, sähköpostiviesti kirjoittajalle 19.12.2018).

Multiplex-RT-PCR sekä siRNA-syväsekvensoinnissa käytetyt Velvet- ja VirusDetect-ohjelmat vahvistivat kokeen kasviaineistossa esiintyvän LuMV:ta, TRV:ta ja ArMV:ta. Kaikki käytetyt menetelmät viittasivat kasviaineistossa esiintyvän potyvirusukuun kuuluvaa lajia, jolle löydettiin tietokannasta vain osittaisia vastaavuuksia. Muista menetelmistä poiketen VirusDetectin tulokset antoivat osittaisia vastaavuuksia hordei- ja closterovirusiin, joskin identtisyysprosentti kontigien ja hordei- ja closterovirusreferenssisekvenssien välillä oli alhainen. Bowtie-tulokset eivät antaneet viitteitä hordei- tai closterovirusten esiintymisestä poolatussa näytteessä, sillä kyseisten referenssisekvenssien peittävyys ja peittävyyden syvyyden arvot jäivät alhaisiksi. Velvet antoi muista menetelmistä poiketen osittaisia vastaavuuksia GFLV:een, mutta koko genomin sijaan ohjelman tulokset löysivät vastaavuuksia vain kontigien ja GFLV:n RNA2-referenssisekvenssin välille. Multiplex-RT-PCR:ssä käytetyt alukkeet eivät ole suunniteltu tunnistamaan hordei- tai closterovirusukuun kuuluvia viruksia, mistä johtuen multiplex-PCR-testien tuloksia ei voida käyttää apuna arvioidessa kyseisten virusten esiintymistä näytteissä. Multiplex A-testissä käytettiin nepovirusten alaryhmä a:ta tunnistavia alukkeita. Tästä johtuen GFLV olisi pitänyt havaita myös multiplex-PCR-testillä, jos kyseistä virusta esiintyi näytteessä.

Poolatusta näytteestä Velvetillä ja VirusDetectillä sekä lupiinikasvinäytteestä JT12 multiplex-RT-PCR:llä löydetty LuMV on potyvirusukuun kuuluva laji, joka löydettiin ensimmäisen kerran vuonna 2009 Tšekistä lupiinikasveista (Sarkisova ja Petrzik 2011). Tulosten perusteella (taulukko 14) eniten LuMV-isolaattia ja näytettä JT12 muistuttava potyviruslaji on luumun rokkovirus (68 %), mikä vastaa Sarkisovan ja Petrzikin tekemää havaintoa LuMV:ta eniten muistuttavista potyviruslajeista.

Tulosten tulkinnassa on huomioitava, että identiteettimatriisien tuloksia ja isolaattien sijoittumista fylogeneettisissä puissa voidaan pitää suuntaa antavana, sillä ne laadittiin koko polyproteiinia koodaavan sekvenssialueen sijaan vain tietyltä genomin alueelta. Fylogeneettisessä puussa (kuva 17) monet bootstrap-arvoista ovat alle 50 %, minkä vuoksi kaikkien potyvirusten sijoittumista puussa ei voida pitää täysin luotettavana.

Poolatusta näytteestä Velvetillä ja VirusDetectillä sekä multiplex-PCR:llä valkome-sikkänäytteestä JT15 löydetylle potyvirusukuun viittaavalle lajille ei löydetty tutkimuksessa NCBI:n tietokannasta täsmällisiä osumia. Nukleotidiyhtäläisyyksien vertailu (taulukko 14) osoitti virusnäytteen JT15 vastaavaan nukleotidisekvenssiltään eniten JYMV:ta (66,4 %) ja toiseksi eniten BYMV:ta (62,8 %). Fylogeneettisessä puussa (kuva 17) näyte JT15 muodosti sukuhaaran BYMV-referenssisekvenssin

kanssa, mutta bootstrap-arvon perusteella (60 %) sukuhaaran sijoittumista fylogeneettisessä puussa ei voida pitää varmana.

Virussekvenssi luokitellaan potyviruslajiksi, mikäli kokonaista polyproteiinia koodaava nukleotidisekvenssi vastaa vähintään 76 %:sti muuta potyviruslajia. Aluetta vastaavalle proteiinisekvenssille rajana pidetään 82 %. Huomioitavaa on, että potyvirusten lajiluokituksen kriteerit vaihtelevat eri genomin alueilla, vaihdellen P1-proteiinia koodaavaan alueen 58 %:sta 74–78 %:iin muille genomin alueille (Adams ym. 2005). Kyseiset lajiluokituksen kriteerit eivät ole ehdottomia, mistä osoituksena on Sarkisovan ja Petrzikin (2011) tekemä tutkimus. Kyseisessä tutkimuksessa kokonaisista polyproteiinisekvensseistä tehty vertailu osoitti nykyään potyvirusiin luokiteltavan LuMV-isolaatin vastanneen 63:sta eri potyviruslajista aminohapposekvenssiltään enintään 54-prosenttisesti muuta potyviruslajia.

On mahdollista, että näytteestä JT15 löydetty sekvenssi kuuluu tuntemattomalle potyviruslajille, joka poikkeaa kohtalaisesti tietokantaan tallennetuista potyvirussekvensseistä. Tämän selvittäminen vaatii jatkotutkimusta, minkä lähtökohtana on selvittää viruksen koko genomi kuten Massart ym. (2017) artikkelissaan esittävät. Mahdollinen löytö uudesta potyviruslajista vahvistaa Zhengin ym. (2010) toteamusta NlB-alukkeiden kyvystä tunnistaa luotettavasti potyviruslajin kuuluvia viruksia.

Poolatusta näytteestä Velvetillä ja VirusDetectillä sekä suikeroalpikasvinäytteestä JT2 multiplex-PCR:llä löydetty TRV on tobravirussukuun kuuluva laji, jota on todettu esiintyvän maailmanlaajuisesti (MacFarlane 1999). Viruksen laajaan isäntäkirjoon kuuluvat monet koriste-, vihannes- ja luonnonvaraiset kasvilajit (MacFarlane 1999, Mojtahedi ym. 2003). Tutkimustulokset vahvistivat suikeroalpien kuuluvan myös osaksi TRV:n isäntäkirjoja.

Bowtie-tulosten perusteella (kuva 13) TRV:n RNA2-referenssisekvenssistä kokonaan ilman peittävyttä jäivät osat 2b- ja 2c-proteiinia koodaavista alueista. Hernándezin ym. (1997) mukaan 2b-proteiinia koodaavaa aluetta vaaditaan nematodivälitteisessä virusinfektiossa. Jälkimmäisen, 2b-proteiinia koodaavan alueen toimintaa ei vielä tiedetä, mutta kyseisen alueen on arvioitu liittyvän myös nematodivälitteiseen virusinfektioon (Ashfaq ym. 2011, Koenig ym. 2015). Tobravirusten RNA2:n koko on todettu vaihtelevan 1900–3900 nukleotidin välillä, riippuen sisältääkö RNA2 proteiiniukuorta koodaavan alueen lisäksi 2b- ja 2c-proteiinia koodaavat alueet vai ei (Ashfaq ym. 2011). Tobraviruksista tiedetään esiintyvän

myös isolaatteja, joilta RNA2 puuttuu kokonaan (MacFarlane 1999). Koenigin ym. (2015) mukaan isäntäkasvissa esiintyvä tobnavirus voi menettää genomistaan 2b- ja 2c-proteiinia koodaavat alueet, mikäli nematodivälitteisestä virusinfektiosta on kulunut riittävän kauan aikaa. Edellä mainittu selitys on mahdollinen syy sille, miksi sRNA-sekvenssit eivät peittäneet kokonaan TRV:n RNA2-referenssisekvenssiä. Koenig ym. (2012) toteavat tobnavirusten RNA2:ssa tapahtuvien deleetioiden mahdollisesti edistävän viruksen sopeutumista eri infektiiovaiheiden olosuhteisiin. Tutkimuksessa löydetty virusisolaatti JT2 vastasi eri TRV-isolaateista eniten Yhdysvalloista löydettyjä perunasta eristettyä isolaattia GQ903771.1 ja pionista (*Paeonia* spp. L.) eristettyä isolaattia JX144383.1 (taulukko 15), jotka muodostivat fylogeneettisessä sukupuussa (kuva 18) luotettavasti (bootstrap-arvo 100 %) oman sukuhaaran.

Poolatusta näytteestä Velvetillä ja VirusDetectillä sekä multiplex-PCR:llä kiinanja-loangervo- ja kvinoakasvinäytteistä 787, 788 ja 790 löydetty ArMV on nepovirusten alaryhmä a:han luokiteltava kasvivirus (Wetzel ym. 2001), jonka esiintymistä kyseisissä kasvinäytteissä epäiltiin perustuen ennen tutkimuksen aloittamista tehtyihin testikasvitesteihin. Bootstrap-arvojen perusteella kaikki ArMV-isolaatit eivät sijoitu fylogeneettisessä puussa luotettavasti, mutta näytteiden muodostamaa sukuhaaraa (bootstrap-arvo 100 %) voidaan pitää luotettavana (kuva 19).

Velvet-ohjelman tuloksista tehty BLAST-haku löysi kontigeille osittaisia vastavuuksia GFLV:hen, mutta koko genomien sijaan osumat vastasivat vain RNA2:ta. GFLV luokitellaan nepovirusten alaryhmä a:han ja on läheistä sukua ArMV:lle (Vigne ym. 2008). Bowtie-tulosten perusteella GFLV:n referenssigenomin peittävyys ja peittävyyden syvyys olivat parhaimpia RNA2:n 5'-pään alussa, jossa esiintyvät UTR-alue ja 2A-proteiinia koodaava alue (kuva 16). Useat tutkimukset ovat osoittaneet, että eri viruslajien välistä rekombinaatiota on tapahtunut ArMV:n ja GFLV:n isolaattien välillä RNA2:n 5'-pään UTR-alueella ja 2A-proteiinia koodaavalla alueella (Lamprecht ym. 2012, Mekuria ym. 2009, Pompe-Novak ym. 2007, Vigne ym. 2008). Vignen ym. (2008) mukaan ArMV:n ja GFLV:n välisen rekombinaation on mahdollistanut niiden esiintyminen samassa isäntäkasvissa useita vuosia, sekvenssien samankaltaisuus ja replikaation aikana tapahtuneet virheet. Tämä selittää miksi jotkin ArMV-isolaateista vastaavat joidenkin genomien alueen osilta enemmän GFLV-isolaatteja ja päinvastoin. Kirjallisuuden perusteella voidaan päätellä, ettei GFLV:ta todennäköisesti esiinny kasvinäytteissä, mutta niistä löydetty ArMV-isolaatti vastaa läheisesti RNA2:n 5'-pään UTR-alueen ja 2A-proteiinia koodaavan alueen osilta joitakin GFLV-isolaatteja.

Velvet-ohjelman tuloksista tehty BLAST-haku löysi kontigeille osittaisia vastaavuuksia GFLV:n lisäksi SMV:hen ja CIYVV:hen. Bowtie-tulosten perusteella sRNA-sekvenssit peittivät vain osittaisesti SMV:n ja CIYVV:n referenssisekvenssejä (kuva 16). Peittävyys ja peittävyiden syvyys osuivat lähinnä genomien alueille, jotka koodaavat Nlb- tai CI-proteiinia (Liu ym. 2016, Sekiguchi ym. 2003). Referenssigenomin 5'-päästä katsottuna peittävyiden syvyydet osuivat pääasiassa nukleotidin sijainneille n. 4500 ja n. 7500–8500 (kuva 16). Kirjallisuuden perusteella BLAST-haun vastaavuudet kyseisiin viruksiin voidaan katsoa selittyvän peittävyiden ja peittävyiden syvyyden osumisella niille genomien alueille, jotka Nicolasin ym. (1991) ja Zhengin ym. (2008) mukaan ovat potyviruslajin sisällä konservoituneita. Käytännössä tämä tarkoittaa, ettei kyseisiä viruksia näytteissä todennäköisesti esiinny.

Illumina-laitteella tehty syväsekvensointi voitiin katsoa onnistuneeksi, sillä sekvensseistä yli 95 % ylitti laatuarvon 30 (0,1 %:n virhetodennäköisyys). Vajaa puolet (46,2 %) sekvensseistä oli pituudeltaan 1–17 nukleotidin mittaisia, mikä viittaa RNA:n osittaiseen hajoamiseen. Mahdollisia syitä RNA:n hajoamiselle ovat etteivät kerätyt kasvinäytteet ole välittömästi keräämisen jälkeen siirretty kylmiöön, RNA:n eristämisen aikana tapahtuneet virheet tai ettei käytetty RNA:n eristämismenetelmä ollut optimi kaikille kasvinäytteille.

Velvet- ja VirusDetect-ohjelmien soveltuvuutta kokoamaan kontigeja ja tunnistamaan siten kasviviruslajia ei voida täysin verrata toisiinsa, sillä VirusDetect tekee Velvetistä poiketen useita eri analyyskejä eikä Velvetin tekemä analyysi kerro esimerkiksi sekvensoinnin syvyydestä. Toisaalta tämän tutkimuksen tulosten perusteella VirusDetectin pisimmät kokoamat kontigit olivat pidempiä kuin Velvetin (taulukot 9 ja 11). Velvetissä pisin muodostettu kontigi oli 1034 nukleotidia pitkä, kun taas VirusDetect tuloksissa LuMV:tä vastannut pisin kontigi oli 6120 nukleotidia pitkä, TRV:n RNA1:tä 3323 nukleotidia ja TRV:n RNA2:tä 1775 nukleotidia.

Viruksista peräisin olevat siRNA-sekvenssit voivat kattaa viruksen koko genomien, kun taas kasvista peräisin olevat, siRNA-sekvenssien pituutta vastaavat pikku-RNA-molekyylit esiintyvät kasvigenomissa epätasaisesti. Tästä johtuen viruksista peräisin olevista siRNA-sekvensseistä kootut kontigit ovat yleensä pidempiä kuin kasvin pikku-RNA-molekyyleistä kootut kontigit (Zheng ym. 2017). Zheng ym. (2017) toteavat edellä mainitusta syystä johtuen parhaimmiksi *De novo* kontigien kokoamisohjelmiksi ne ohjelmat, jotka pystyvät kokoamaan pisimpiä kontigeja. Väittämän perusteella tämän tutkimuksen tulosten voidaan katsoa puoltavan enemmän VirusDetectin kuin Velvetin käyttöä kasviviruslajien tunnistamisessa.

7 Johtopäätökset

Tutkimuksessa pystyttiin luomaan kaksi multiplex-RT-PCR-testiä, jotka tunnistavat yhteensä kahdeksaan eri virusryhmään kuuluvia viruksia. Saadut tulokset toimivat lähtökohtana multiplex-RT-PCR-testien validoinnille, jonka avulla voidaan varmistua testien antamien tulosten luotettavuudesta ja soveltuvuudesta käyttötarkoitukseensa.

Tulokset osoittivat siRNA-syväseksensoinnin tunnistaneen osasta löytyneistä viruksista lähes niiden koko genomin. VirusDetect- ja Velvet-ohjelmien kokoamien kontigien vertailu osoitti VirusDetectin koonneen pidempiä kontigeja siRNA-sekvensseistä kuin Velvet.

Työssä kartoitetuista kasvinäytteistä optimoidut multiplex-RT-PCR-testit havaitsivat lupiinin mosaiikkiviruksen (LuMV) lupiinista (*Lupinus polyphyllus* Lindl.), arabiksen mosaiikkiviruksen (ArMV) kiinanjaloangervosta (*Astilbe chinensis* (Maxim.) Franch.) ja kvinoasta (*Chenopodium quinoa* Willd.), tupakan rattleviruksen (TRV) suikeroalpista (*Lysimachia nummularia* L.) ja mahdollisesti uuden potyviruslajin valkomesikästä (*Melilotus albus* Medik.). siRNA-syväseksensointi tunnisti samat kasvivirukset näytteistä kuin mitä multiplex-RT-PCR-testit, mikä vahvistaa luottamusta optimoituja multiplex-RT-PCR-testejä kohtaan.

8 Kiitokset

Haluan esittää erityiset kiitokset ohjaajilleni Johanna Santalalle, Minna Rajamäelle ja Jari Valkoselle neuvoista ja ohjauksesta sekä kokeiden toteutuksessa että varsinaisessa kirjoitusprosessissa. Kiitokset Juha Tuomolalle luonnonkasvinäytteiden keruusta. Kiitokset kasvintuhoojalaboratorion henkilökunnalle käytännön neuvoista laboratoriotyöskentelyyn. Lopuksi haluan kiittää Jimi Juolaa korvaamattomasta tuesta koko prosessin ajalta.

Lähteet

- Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navaliniskiene, M., Samuitiene, M. & Boonham, N. 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular Plant Pathology* 10 (4): 537–545.
- Adams, I. & Fox, A. 2016. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis. *Teoksessa: Current research topics in plant virology*. Toim. Wang, A. & Zhou, X. Springer International Publishing. S. 323–335.
- Adams, M. J., Antoniw, J. F. & Fauquet, C. M. 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Archives of Virology* 150 (3): 459–479.
- Agrios, G. N. 2005. *Plant Pathology*. 5th. Elsevier. 947 s.
- Al Rwahnih, M., Daubert, S., Golino, D. & Rowhani, A. 2009. Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology* 387 (2): 395–401.
- Altschul, S., Gish, W., Miller, W., Myers, E. & Lipman, D. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215 (3): 403–410.
- Ashfaq, M., McGavin, W. & MacFarlane, S. 2011. RNA2 of TRV SYM breaks the rules for tobnavirus genome structure. *Virus Research* 160 (1-2): 435–438.
- Baulcombe, D. 2004. RNA silencing in plants. *Nature* 431 (7006): 356–363.
- Baumberger, N., Tsai, C.-H., Lie, M., Havecker, E. & Baulcombe, D. C. 2007. The Polerovirus silencing suppressor P0 Targets ARGONAUTE proteins for degradation. *Current Biology* 17 (18): 1609–1614.
- Blevins, T., Rajeswaran, R., Shivaprasad, P. V., Becnazariants, D., Si-Ammour, A., Park, H.-S., Vazquez, F., Robertson, D., Meins, F., Hohn, T. & Pooggin, M. M. 2006. Four plant Dicers mediate viral small RNA biogenesis and DNA virus induced silencing. *Nucleic Acids Research* 34 (21): 6233–6246.
- Boonham, N., Kreuze, J., Winter, S., van der Vlugt, R., Bergervoet, J., Tomlinson, J. & Mumford, R. 2014. Methods in virus diagnostics: From ELISA to next generation sequencing. *Virus Research. Plant Virus Epidemiology Symposium Ecology, Evolution and Control of Plant Viruses and Their Vectors* 186: 20–31.

- Candresse, T., Filloux, D., Muhire, B., Julian, C., Galzi, S., Fort, G., Bernardo, P., Daugrois, J.-H., Fernandez, E., Martin, D. P., Varsani, A. & Roumagnac, P. 2014. Appearances can be deceptive: revealing a hidden viral infection with deep sequencing in a plant quarantine context. *PLOS ONE* 9 (7): e102945.
- Carbonell, A. & Carrington, J. C. 2015. Antiviral roles of plant ARGONAUTES. *Current Opinion in Plant Biology. Cell Signalling and Gene Regulation* 27: 111–117.
- Chan, S. W.-L., Zilberman, D., Xie, Z., Johansen, L. K., Carrington, J. C. & Jacobsen, S. E. 2004. RNA silencing genes control *de novo* DNA methylation. *Science* 303 (5662): 1336–1336.
- Chang, S., Puryear, J. & Cairney, J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11 (2): 113–116.
- Chen, T.-C., Li, J.-T., Lin, Y.-P., Yeh, Y.-C., Kang, Y.-C., Huang, L.-H. & Yeh, S.-D. 2012. Genomic characterization of *Calla lily chlorotic spot virus* and design of broad-spectrum primers for detection of tospoviruses. *Plant Pathology* 61 (1): 183–194.
- Chikhi, R. & Medvedev, P. 2014. Informed and automated k-mer size selection for genome assembly. *Bioinformatics* 30 (1): 31–37.
- Clark, M. F. & Adams, A. N. 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *The Journal of General Virology* 34 (3): 475–483.
- Coetzee, B., Freeborough, M.-J., Maree, H. J., Celton, J.-M., Rees, D. J. G. & Burger, J. T. 2010. Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard. *Virology* 400 (2): 157–163.
- Cooper, I. & Jones, R. A. C. 2006. Wild plants and viruses: under-investigated ecosystems. *Teoksessa: Plant Virus Epidemiology. Toim. Thresh, J. M. Vol. 67. San Diego: Elsevier Academic Press Inc. S. 1–47.*
- Cornishbowden, A. 1985. Nomenclature for incompletely specified bases in nucleic-acid sequences -recommendations 1984. *Nucleic Acids Research* 13 (9): 3021–3030.
- Dietzgen, R. G. 2002. Application of PCR in plant virology. *Teoksessa: Khan, J. A. & Dijkstra, J. Plant viruses as molecular pathogens. New York: Food Products Press. s. 471–500.*
- Digiario, M., Elbeaino, T. & Martelli, G. P. 2007. Development of degenerate and species-specific primers for the differential and simultaneous RT-PCR detection of

grapevine-infecting nepoviruses of subgroups A, B and C. *Journal of Virological Methods* 141 (1): 34–40.

Ding, S.-W. & Voinnet, O. 2007. Antiviral immunity directed by small RNAs. *Cell* 130 (3): 413–426.

Efron, B. 1979. 1977 Rietz Lecture - bootstrap methods - another look at the jackknife. *Annals of Statistics* 7 (1): 1–26.

Elnifro, E. M., Ashshi, A. M., Cooper, R. J. & Klapper, P. E. 2000. Multiplex PCR: optimization and application in diagnostic virology. *Clinical Microbiology Reviews* 13 (4): 559–570.

EPPO. 2004. Potato spindle tuber pospiviroid. *EPPO Bulletin* 34 (2): 257–269.

EPPO. 2014. PM 7/98 (2) Specific requirements for laboratories preparing accreditation for a plant pest diagnostic activity. *EPPO Bulletin* 44 (2): 117–147.

Euroopan unionin neuvosto. 2000. Neuvoston direktiivi 2000/29/EY, annettu 8. päivänä toukokuuta 2000, kasvien ja kasvituotteiden haitallisten organismien jäsenvaltioihin kulkeutumisen estämiseen liittyvistä suojatoimenpiteistä.

Evira. 2018. Kasvinterveyden, taimiaineiston ja metsänviljelyaineiston valvonnan raportti 2017, Osa I Kasvinterveys ja taimiaineisto. 2093/0411/2018. S. 25.

FAO. 2016. Framework for pest risk analysis. *International Standards for Phytosanitary Measures* (02).

FAO. 2017a. Pest risk analysis for quarantine pests. *International Standards for Phytosanitary Measures* (11).

FAO. 2017b. The future of food and agriculture -trends and challenges. Rome: Food and Agriculture Organization of the United Nations. 163 s.

Ferragina, P. & Manzini, G. 2000. Opportunistic data structures with applications. Teoksessa: 41st Annual Symposium on Foundations of Computer Science, Proceedings. Los Alamitos: IEEE Computer Society. S. 390–398.

Fraenkel-Conrat, H. 1956. The role of the nucleic acid in the reconstitution of active *Tobacco mosaic virus*. *Journal of the American Chemical Society* 78 (4): 882–883.

Gambino, G. & Gribaudo, I. 2006. Simultaneous detection of nine grapevine viruses by multiplex reverse transcription-polymerase chain reaction with coamplification of a plant RNA as internal control. *Phytopathology* 96 (11): 1223–1229.

- Gladman, S. & Seemann, T. 2012. Victorian Bioinformatics Consortium - VelvetOptimiser. URL: <http://www.vicbioinformatics.com/software.velvetoptimiser.shtml> (viitattu 4. toukokuuta 2018).
- Goelet, P., Lomonossoff, G. P., Butler, P. J., Akam, M. E., Gait, M. J. & Karn, J. 1982. Nucleotide sequence of *Tobacco mosaic virus* RNA. Proceedings of the National Academy of Sciences of the United States of America 79 (19): 5818–5822.
- Grieco, F. & Gallitelli, D. 1999. Multiplex reverse transcriptase-polymerase chain reaction applied to virus detection in globe artichoke. Journal of Phytopathology 147 (3): 183–185.
- Grunenwald, H. 2003. Optimization of polymerase chain reactions. Teoksessa: PCR Protocols. Toim. Bartlett, J. M. S. & David Stirling. Methods in Molecular Biology™. Humana Press. S. 89–99.
- Hadidi, A., Flores, R., Candresse, T. & Barba, M. 2016. Next-generation sequencing and genome editing in plant virology. Frontiers in Microbiology 7.
- Hamilton, A. J. & Baulcombe, D. C. 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science 286 (5441): 950–952.
- Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H. & Vogt, P. H. 1997. Multiplex PCR: critical parameters and step-by-step protocol. Biotechniques 23 (3): 504–511.
- Henson, J. M. & French, R. 1993. The polymerase chain reaction and plant disease diagnosis. Annual Review of Phytopathology 31 (1): 81–109.
- Hernández, C., Visser, P., Brown, D. & Bol, J. 1997. Transmission of tobacco rattle virus isolate PpK20 by its nematode vector requires one of the two non-structural genes in the viral RNA 2. Journal of General Virology 78 (2): 465–467.
- Hu, C.-C., Hsu, Y.-H. & Lin, N.-S. 2009. Satellite RNAs and satellite viruses of plants. Viruses 1 (3): 1325–1350.
- Hull, R. 2014. Plant virology. Fifth edition. Amsterdam ; Boston: Elsevier/AP, Academic Press is an imprint of Elsevier. 1104 s.
- Hyndman, D. L. & Mitsuhashi, M. 2003. PCR primer design. Teoksessa: PCR Protocols. Toim. Bartlett, J. M. S. & Stirling, D. Methods in Molecular Biology. Humana Press. S. 81–88.
- James, D., Varga, A., Pallas, V. & Candresse, T. 2006. Strategies for simultaneous detection of multiple plant viruses. Canadian Journal of Plant Pathology 28 (1): 16–29.

- James, D. 1999. A simple and reliable protocol for the detection of apple stem grooving virus by RT-PCR and in a multiplex PCR assay. *Journal of Virological Methods* 83 (1): 1–9.
- Jo, Y., Choi, H., Kim, S.-M., Kim, S.-L., Lee, B. C. & Cho, W. K. 2016. Integrated analyses using RNA-Seq data reveal viral genomes, single nucleotide variations, the phylogenetic relationship, and recombination for *Apple stem grooving virus*. *Bmc Genomics* 17: 579.
- Jones, D., Farreyrol, K., Clover, G. R. G. & Pearson, M. N. 2008. Development of a generic PCR detection method for tobnaviruses. *Australasian Plant Pathology* 37 (2): 132–136.
- Jones, R. A. C. 2009. Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Research. Plant Virus Epidemiology: Controlling Epidemics of Emerging and Established Plant Viruses - the Way Forward* 141 (2): 113–130.
- Jones, R. A. C. 2014. Trends in plant virus epidemiology: opportunities from new or improved technologies. *Virus Research. Plant Virus Epidemiology Symposium Ecology, Evolution and Control of Plant Viruses and Their Vectors* 186: 3–19.
- Jones, S., Baizan-Edge, A., MacFarlane, S. & Torrance, L. 2017. Viral diagnostics in plants using next generation sequencing: computational analysis in practice. *Frontiers in Plant Science* 8.
- Kansainvälinen virusluokituskomitea. 2017. Genus: Nepovirus. URL: https://talk.ictvonline.org/ictv-reports/ictv_online_report/positive-sense-rna-viruses/picornavirales/w/secoviridae/591/genus-nepovirus (viitattu 1. kesäkuuta 2018).
- Kashif, M., Pietila, S., Artola, K., Jones, R. A. C., Tugume, A. K., Makinen, V. & Valkonen, J. P. T. 2012. Detection of viruses in sweetpotato from Honduras and Guatemala augmented by deep-sequencing of small-RNAs. *Plant Disease* 96 (10): 1430–1437.
- Kimura, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16 (2): 111–120.
- Koenig, R., Lesemann, D.-E. & Pleij, C. W. A. 2012. *Tobacco Rattle Virus* genome alterations in the *Hosta* hybrid ‘Green Fountain’ and other plants: reassortments, recombinations and deletions. *Archives of Virology* 157 (10): 2005–2008.

- Koenig, R., Hilbrich, I. & Lindner, K. 2015. Complete nucleotide sequences of two closely related tobacco rattle virus RNA2s for which previously possibly only deletion/recombination mutants had been described. *Virus Research* 203: 20–23.
- Kolchinsky, A., Kolesnikova, M. & Ananiev, E. 1991. Portraying of plant genomes using polymerase chain-reaction amplification of ribosomal 5s genes. *Genome* 34 (6): 1028–1031.
- Kreuze, J. 2014. siRNA deep sequencing and assembly: piecing together viral infections. *Teoksessa: Detection and diagnostics of plant pathogens*. Toim. Aiming, W. & Xueping, Z. *Plant Pathology in the 21st Century*. Springer, Dordrecht. S. 21–38.
- Kreuze, J., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I. & Simon, R. 2009. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388 (1): 1–7.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35 (6). Toim. Battistuzzi, F. U.: 1547–1549.
- Kutnjak, D., Silvestre, R., Cuellar, W., Perez, W., Müller, G., Ravnikar, M. & Kreuze, J. 2014. Complete genome sequences of new divergent *potato virus X* isolates and discrimination between strains in a mixed infection using small RNAs sequencing approach. *Virus Research* 191: 45–50.
- Lakatos, L., Csorba, T., Pantaleo, V., Chapman, E. J., Carrington, J. C., Liu, Y.-P., Dolja, V. V., Calvino, L. F., López-Moya, J. J. & Burgyán, J. 2006. Small RNA binding is a common strategy to suppress RNA silencing by several viral suppressors. *The EMBO Journal* 25 (12): 2768–2780.
- Lamprecht, R. L., Maree, H. J., Stephan, D. & Burger, J. T. 2012. Complete nucleotide sequence of a South African isolate of *Grapevine fanleaf virus*. *Virus Genes* 45 (2): 406–410.
- Langham, M. 2007. Plant pathogenic viruses. *Teoksessa: Plant pathology -concepts and laboratory exercises*. Toim. Trigiano, R., Windham, M. & Windham, A. 2nd. CRC Press. S. 21–34.
- Langham, M. A. & Brown, J. K. 2016. Plant pathogenic viruses. *Teoksessa: Plant pathology -concepts and laboratory exercises*. Toim. Ownley, B. H. & Trigiano, R. N. 3. painos. Boca Raton: CRC Press. S. 51–80.

- Langmead, B., Daehwan, K., Charles, R., Wilks, C. & Antonescu, V. 2017. Bowtie 2: fast and sensitive read alignment. URL: <http://bowtie-bio.sourceforge.net/bowtie2/index.shtml> (viitattu 11. heinäkuuta 2017).
- Langmead, B. & Salzberg, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9 (4): 357–U54.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23 (21): 2947–2948.
- Lemmetty, A., Laamanen, J., Soukainen, M. & Tegel, J. 2011. Emerging virus and viroid pathogen species identified for the first time in horticultural plants in Finland in 1997-2010. *Agricultural and Food Science* 20 (1): 29–41.
- Letscher, B., Adam, G., Lesemann, D. E., Willingham, P. & Heinze, C. 2002. Detection and differentiation of serologically cross-reacting tobamoviruses of economical importance by RT-PCR and RT-PCR-RFLP. *Journal of Virological Methods* 106 (1): 1–10.
- Li, H. & Durbin, R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25 (14): 1754–1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. & Durbin, R. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25 (16): 2078–2079.
- Li, H., Ruan, J. & Durbin, R. 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Research* 18 (11): 1851–1858.
- Li, R., Gao, S., Hernandez, A. G., Wechter, W. P., Fei, Z. & Ling, K.-S. 2012. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLOS ONE* 7 (5).
- Linhart, C. & Shamir, R. 2005. The degenerate primer design problem: theory and applications. *Journal of Computational Biology* 12 (4): 431–456.
- Liu, J.-Z., Fang, Y. & Pang, H. 2016. The current status of the Soybean-Soybean Mosaic Virus (SMV) pathosystem. *Frontiers in Microbiology* 7 (NOV).
- Lomonosoff, G. P. 2018. So what have plant viruses ever done for virology and molecular biology? Teoksessa: *Advances in Virus Research*. Toim. Kielian, M., Mettenleiter, T. C. & Roossinck, M. J. Vol. 100. Academic Press. S. 145–162.

- López, M. M., Bertolini, E., Olmos, A., Caruso, P., Gorris, M. T., Llop, P., Penyalver, R. & Cambra, M. 2003. Innovative tools for detection of plant pathogenic viruses and bacteria. *International Microbiology: The Official Journal of the Spanish Society for Microbiology* 6 (4): 233–243.
- Luck, J., Spackman, M., Freeman, A., TreBicki, P., Griffiths, W., Finlay, K. & Chakraborty, S. 2011. Climate change and diseases of food crops. *Plant Pathology* 60 (1): 113–121.
- Maa- ja metsätalousministeriö. 2008. Asetus kasvinterveyden suojelemisesta, nro 17/08.
- MacFarlane, S. A. 1999. Molecular biology of the tobnaviruses. *Journal of General Virology* 80: 2799–2807.
- Mackay, I. M., Arden, K. E. & Nitsche, A. 2002. Real-time PCR in virology. *Nucleic Acids Research* 30 (6): 1292–1305.
- MacKenzie, D. J., McLean, M. A., Mukerji, S. & Green, M. 1997. Improved RNA extraction from woody plants for the detection of viral pathogens by reverse transcription-polymerase chain reaction. *Plant Disease* 81 (2): 222–226.
- Markoulatos, P., Siafakas, N. & Moncany, M. 2002. Multiplex polymerase chain reaction: a practical approach. *Journal of Clinical Laboratory Analysis* 16 (1): 47–51.
- Martin, R. R., James, D. & Lévesque, C. A. 2000. Impacts of molecular diagnostic technologies on plant disease management. *Annual Review of Phytopathology* 38 (1): 207–239.
- Massart, S., Candresse, T., Gil, J., Lacomme, C., Predajna, L., Ravnikar, M., Reynard, J.-S., Rumbou, A., Saldarelli, P., Škorić, D., Vainio, E. J., Valkonen, J. P. T., Vanderschuren, H., Varveri, C. & Wetzels, T. 2017. A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Frontiers in Microbiology* 8.
- Massart, S., Olmos, A., Jijakli, H. & Candresse, T. 2014. Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research* 188: 90–96.
- McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. & Lopez, R. 2013. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Research* 41 (W1): W597–W600.

- Mekuria, T. A., Gutha, L. R., Martin, R. R. & Naidu, R. A. 2009. Genome diversity and intra- and interspecies recombination events in *Grapevine fanleaf virus*. *Phytopathology* 99 (12): 1394–1402.
- Menzel, W., Jelkmann, W. & Maiss, E. 2002. Detection of four apple viruses by multiplex RT-PCR assays with coamplification of plant mRNA as internal control. *Journal of Virological Methods* 99 (1-2): 81–92.
- Metzker, M. L. 2010. Sequencing technologies – the next generation. *Nature Reviews Genetics* 11 (1): 31–46.
- Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D. & Marshall, D. 2013. Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics* 14 (2): 193–202.
- Mlotshwa, S., Pruss, G. J. & Vance, V. 2008. Small RNAs in viral infection and host defense. *Trends in Plant Science* 13 (7): 375–382.
- Mojtahedi, H., Boydston, R. A., Thomas, R. E., Crosslin, J. M., Santo, G. S., Riga, E. & Anderson, I. L. 2003. Weed hosts of *Paratrichodorus allius* and *Tobacco rattle virus* in the pacific northwest. *American Journal of Potato Research* 80 (6): 379–385.
- Morozov, S. Y., Ryabov, E. V., Leiser, R.-M. & Zavriev, S. K. 1995. Use of highly conserved motifs in plant virus RNA polymerases as the tags for specific detection of carmovirus-related RNA-dependent RNA polymerase genes. *Virology* 207 (1): 312–315.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology* 51 Pt 1: 263–273.
- Mullis, K. 1990. The unusual origin of the polymerase chain reaction. *Scientific American* 262 (4): 56–65.
- Narayanasamy, P. 2010. Microbial plant pathogens - detection and disease diagnosis: bacterial and phytoplasmal pathogens. Springer Science & Business Media. 277 s.
- Nassuth, A., Pollari, E., Helmeczy, K., Stewart, S. & Kofalvi, S. A. 2000. Improved RNA extraction and one-tube RT-PCR assay for simultaneous detection of control plant RNA plus several viruses in plant extracts. *Journal of Virological Methods* 90 (1): 37–49.
- Nicolas, O. & Laliberté, J.-F. 1991. The use of PCR for cloning of large cDNA fragments of *Turnip mosaic potyvirus*. *Journal of Virological Methods* 32 (1): 57–66.

Nonomura, K.-I., Morohoshi, A., Nakano, M., Eiguchi, M., Miyao, A., Hirochika, H. & Kurata, N. 2007. A germ cell-specific gene of the ARGONAUTE family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. *Plant Cell* 19 (8): 2583–2594.

Novocraft Technologies Sdn Bhd. 2014. NovoAlign | Novocraft. URL: <http://www.novocraft.com/products/novoalign/> (viitattu 11. heinäkuuta 2017).

Olmos, A., Boonham, N., Candresse, T., Gentit, P., Giovani, B., Kutnjak, D., Liefting, L., Maree, H. J., Minafra, A., Moreira, A., Nakhla, M. K., Petter, F., Ravnikar, M., Rodoni, B., Roenhorst, J. W., Rott, M., Ruiz-García, A. B., Santala, J., Stancanelli, G., Vlugt, R. van der, Varveri, C., Westenberg, M., Wetzel, T., Ziebell, H. & Massart, S. 2018. High-throughput sequencing technologies for plant pest diagnosis: challenges and opportunities. *EPPO Bulletin* 48 (2): 219–224.

Pallás, V., Sánchez-Navarro, J. A. & James, D. 2018. Recent advances on the multiplex molecular detection of plant viruses and viroids. *Frontiers in Microbiology* 9.

Pecman, A., Kutnjak, D., Gutiérrez-Aguirre, I., Adams, I., Fox, A., Boonham, N. & Ravnikar, M. 2017. Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches. *Frontiers in Microbiology* 8.

Penczykowski, R. M., Laine, A.-L. & Koskella, B. 2016. Understanding the ecology and evolution of host-parasite interactions across scales. *Evolutionary Applications* 9 (1): 37–52.

Pompe-Novak, M., Gutiérrez-Aguirre, I., Vojvoda, J., Blas, M., Tomažič, I., Vigne, E., Fuchs, M., Ravnikar, M. & Petrovič, N. 2007. Genetic variability within RNA2 of *Grapevine fanleaf virus*. *European Journal of Plant Pathology* 117 (3): 307–312.

Pooggin, M. M. 2018. Small RNA-omics for plant virus identification, virome reconstruction, and antiviral defense characterization. *Frontiers in Microbiology* 9.

Qiagen. 2012. RNeasy® Mini Handbook, Fourth edition.

Rinehart, T., Wang, X. & Trigiano, R. 2007. Molecular tools for studying plant pathogens. *Teoksessa: Plant Pathology Concepts and Laboratory Exercises*. Toim. Trigiano, R., Windham, M. & Windham, A. 2nd. CRC Press. S. 269–279.

Roossinck, M. J. 1997. Mechanisms of plant virus evolution. *Annual Review of Phytopathology* 35 (1): 191–209.

Roossinck, M. J. 2012. Plant virus metagenomics: biodiversity and ecology. *Annual Review of Genetics* 46 (1): 359–369.

- Roossinck, M. J. 2015. A new look at plant viruses and their potential beneficial roles in crops. *Molecular Plant Pathology* 16 (4): 331–333.
- Roossinck, M. J. 2017. Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Research. Deep Sequencing in Virology* 239: 82–86.
- Roossinck, M. J., Martin, D. P. & Roumagnac, P. 2015. Plant virus metagenomics: advances in virus discovery. *Phytopathology* 105 (6): 716–727.
- Roy, A., Ananthakrishnan, G., Hartung, J. S. & Brlansky, R. H. 2010. Development and application of a multiplex reverse-transcription polymerase chain reaction assay for screening a global collection of *Citrus tristeza virus* isolates. *Phytopathology* 100 (10): 1077–1088.
- Saade, M., Aparicio, F., Sanchez-Navarro, J. A., Herranz, M. C., Myrta, A., Di Terlizzi, B. & Pallas, V. 2000. Simultaneous detection of the three ilarviruses affecting stone fruit trees by nonisotopic molecular hybridization and multiplex reverse-transcription polymerase chain reaction. *Phytopathology* 90 (12): 1330–1336.
- Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4 (4): 406–425.
- Sanchez-Navarro, J. A., Aparicio, F., Herranz, M. C., Minafra, A., Myrta, A. & Pallas, V. 2005. Simultaneous detection and identification of eight stone fruit viruses by one-step RT-PCR. *European Journal of Plant Pathology* 111 (1): 77–84.
- Santala, J. & Soukainen, M. 2015. First report of *Plum pox virus* on *plum* in Finland. *EPPO Bulletin* 45 (2): 193–194.
- Santala, J. & Valkonen, J. P. T. 2018. Sensitivity of small RNA-based detection of plant viruses. *Frontiers in Microbiology* 9.
- Sarkisova, T. & Petrzik, K. 2011. Determination of the complete nucleotide sequence of a lupine potyvirus isolate from the Czech Republic reveals that it belongs to a new member of the genus *Potyvirus*. *Archives of Virology* 156 (1): 167–169.
- Sastry, K. S. 2013. Plant virus and viroid diseases in the tropics. Vol. 1. Springer Netherlands. 361 s.
- Savary, S., Ficke, A., Aubertot, J.-N. & Hollier, C. 2012. Crop losses due to diseases and their implications for global food production losses and food security. *Food Security* 4 (4): 519–537.

- Sekiguchi, H., Tacahashi, Y. & Uyeda, I. 2003. The 3' terminal region is strictly required for *Clover yellow vein virus* genome replication. *Archives of Virology* 148 (4): 759–772.
- Sharman, M., Thomas, J. E. & Dietzgen, R. G. 2000. Development of a multiplex immunocapture PCR with colourimetric detection for viruses of banana. *Journal of Virological Methods* 89 (1): 75–88.
- Singh, R. P. 1998. Reverse-transcription polymerase chain reaction for the detection of viruses from plants and aphids. *Journal of Virological Methods* 74 (2): 125–138.
- Soosaar, J. L. M., Burch-Smith, T. M. & Dinesh-Kumar, S. P. 2005. Mechanisms of plant resistance to viruses. *Nature Reviews Microbiology* 3 (10): 789–798.
- Stobbe, A. H. & Roossinck, M. J. 2014. Plant virus metagenomics: what we know and why we need to know more. *Frontiers in Plant Science* 5.
- Stucky, B. J. 2012. SeqTrace: a graphical tool for rapidly processing DNA sequencing chromatograms. *Journal of Biomolecular Techniques : JBT* 23 (3): 90–93.
- Tapio, E., Bremer, K. & Valkonen, J. P. 1997. Viruses and their significance in agricultural and horticultural crops in Finland. *Agricultural and Food Science in Finland* 6 (4): 323–336.
- Tugume, A. K., Mukasa, S. B. & Valkonen, J. P. T. 2008. Natural wild hosts of *Sweet potato feathery mottle virus* show spatial differences in virus incidence and virus-like diseases in Uganda. *Phytopathology* 98 (6): 640–652.
- Tuo, D., Shen, W., Yang, Y., Yan, P., Li, X. & Zhou, P. 2014. Development and validation of a multiplex reverse transcription PCR assay for simultaneous detection of three papaya viruses. *Viruses-Basel* 6 (10): 3893–3906.
- Untiveros, M., Perez-Egusquiza, Z. & Clover, G. 2010. PCR assays for the detection of members of the genus *Ilarvirus* and family *Bromoviridae*. *Journal of Virological Methods* 165 (1): 97–104.
- Valtioneuvosto. 1994. Taimiaineistolaki 1205/1994. Annettu 16.12.1994. URL: <https://www.finlex.fi/fi/laki/ajantasa/1994/19941205> (viitattu 6. toukokuuta 2018).
- Valtioneuvosto. 2003. Laki kasvinterveyden suojelemisesta. Laki n:o 702. Annettu 18.7.2003. Suomen säädöskokoelma N:o 701–707 (Vihko 110): 2816–2826.
- Valtioneuvosto. 2018. Laki Ruokavirastosta. Laki n:o 371. Annettu 18.5.2018. URL: <https://www.finlex.fi/fi/laki/kokoelma/2018/sk20180371.pdf> (viitattu 23. joulukuuta 2018).

- Van Pelt-Verkuil, E., van Belkum, A. & Hays, J. P. 2008. Principles and technical aspects of PCR amplification. Springer Science & Business Media. 333 s.
- Wang, M.-B., Masuta, C., Smith, N. A. & Shimura, H. 2012. RNA silencing and plant viral diseases. *Molecular Plant-Microbe Interactions* 25 (10): 1275–1285.
- Ward, E., Foster, S. J., Fraaije, B. A. & McCartney, H. A. 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Annals of Applied Biology* 145 (1): 1–16.
- Wei, T., Lu, G. & Clover, G. R. G. 2009. A multiplex RT-PCR for the detection of *Potato yellow vein virus*, *Tobacco rattle virus* and *Tomato infectious chlorosis virus* in potato with a plant internal amplification control. *Plant Pathology* 58 (2): 203–209.
- Wei, T. & Clover, G. 2008. Use of primers with 5' non-complementary sequences in RT-PCR for the detection of nepovirus subgroups A and B. *Journal of Virological Methods* 153 (1): 16–21.
- Wetzel, T., Meunier, L., Jaeger, U., Reustle, G. & Krczal, G. 2001. Complete nucleotide sequences of the RNAS 2 of German isolates of grapevine fanleaf and arabis mosaic nepoviruses. *Virus Research* 75 (2): 139–145.
- Whitfield, A. E., Falk, B. W. & Rotenberg, D. 2015. Insect vector-mediated transmission of plant viruses. *Virology*. 60th Anniversary Issue 479-480: 278–289.
- Vigne, E., Marmonier, A. & Fuchs, M. 2008. Multiple interspecies recombination events within RNA2 of *Grapevine fanleaf virus* and *Arabis mosaic virus*. *Archives of Virology* 153 (9): 1771–1776.
- Vlugt, R. A. A. van der & Berendsen, M. 2002. Development of a general *Potexvirus* detection method. *European Journal of Plant Pathology* 108 (4): 367–371.
- Wu, Q., Ding, S.-W., Zhang, Y. & Zhu, S. 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. *Annual Review of Phytopathology* 53 (1): 425–444.
- Wylie, S. J. & Jones, M. G. K. 2011. The complete genome sequence of a *Passion fruit woodiness virus* isolate from Australia determined using deep sequencing, and its relationship to other potyviruses. *Archives of Virology* 156 (3): 479–482.
- Wylie, S. J., Luo, H., Li, H. & Jones, M. G. K. 2012. Multiple polyadenylated RNA viruses detected in pooled cultivated and wild plant samples. *Archives of Virology* 157 (2): 271–284.
- Xu, P., Chen, F., Mannas, J. P., Feldman, T., Sumner, L. W. & Roossinck, M. J. 2008. Virus infection improves drought tolerance. *New Phytologist* 180 (4): 911–921.

- Zerbino, D. R. 2010. Using the velvet *de novo* assembler for short-read sequencing technologies. *Current Protocols in Bioinformatics* 31 (1): 11.5.1–11.5.12.
- Zerbino, D. R. & Birney, E. 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* 18 (5): 821–829.
- Zhang, Y., Singh, K., Kaur, R. & Qiu, W. 2011. Association of a novel DNA virus with the grapevine vein-clearing and vine decline syndrome. *Phytopathology* 101 (9): 1081–1090.
- Zhang, Y., Yu, N., Huang, Q., Yin, G., Guo, A., Wang, X., Xiong, Z. & Liu, Z. 2014. Complete genome of Hainan *Papaya ringspot virus* using small RNA deep sequencing. *Virus Genes* 48 (3): 502–508.
- Zheng, L., Rodoni, B. C., Gibbs, M. J. & Gibbs, A. J. 2010. A novel pair of universal primers for the detection of potyviruses. *Plant Pathology* 59 (2): 211–220.
- Zheng, L., Wayper, P. J., Gibbs, A. J., Fourment, M., Rodoni, B. C. & Gibbs, M. J. 2008. Accumulating variation at conserved sites in potyvirus genomes is driven by species discovery and affects degenerate primer design. *PLoS ONE* 3 (2).
- Zheng, Y., Gao, S., Padmanabhan, C., Li, R., Galvez, M., Gutierrez, D., Fuentes, S., Ling, K.-S., Kreuze, J. & Fei, Z. 2017. VirusDetect: An automated pipeline for efficient virus discovery using deep sequencing of small RNAs. *Virology* 500: 130–138.

Liite A: Vaihtelevat emäket

Tutkimuksessa käytettyjen degeneroitujen alukkeiden sisältämät vaihtelevat emäket sekä niitä vastaavat emäket (Cornishbowden 1985).

Vaihtelevat emäket	Vastaavat emäket
R	A tai G
Y	C tai T
M	A tai C
K	G tai T
S	C tai G
W	A tai T
H	A, C, T
B	C, G tai T
V	A, C tai G
D	A, G tai T
N	A, C, G tai T

Liite B: Multiplex-RT-PCR-sekvenssit

Multiplex-RT-PCR:ssä löydettyjen virussekvenssien vertailu parhaimman BLAST-osuman antaneen virussekvenssin kanssa.

Arabis mosaic virus RNA-dependent RNA polymerase mRNA, partial cds

Sequence ID: EU741688.1

790	GAN	GGG.TAN	CCTGAAGTCCTTCAG	CG	AAAAAGC	AGGAGA.	38
EU741688.1	GAG	GGG.TAT	CCTGAAGTCCTTCAG	AG	AAAAAGC	TGGAGA.	43
790	AAA	GGGT	TAAAGCCCGATT	TTTTGT	TGGGGAGCCTGGAGC	G	78
EU741688.1	AAA	AGG	CAAAGCCCGATT	CTTTGT	AGGGGAGCCTGGAGCA	A	83
790	CGT	GAATTTGTGCC	CGGGTGTGGTCCTGAAAAA	ACTTAT	T		118
EU741688.1	CGG	GAATTTGTGCC	TGGGTGTGGTCCTGAAAAA	GCTTAT	C		123
790	TGAGTCTAGAAGAGGAATGCAAAACGCGTGTCCCTTCTCT						158
EU741688.1	TGAGTCTAGAAGAGGAATGCAAAACGCGTGTCCCTTCTCT						163
790	TATAAGT	TATTGAG	ACCCCAAAGATGAAAGGCT	TAAAAGA	A		198
EU741688.1	TATAAG	CATTGAA	ACCCCAAAGATGAAAGGCT	CAAGAGG	G		203
790	TCT	TAA	GATTGAAACTCCAGGTACGAG	A	TTGTTTTCTGTGC		238
EU741688.1	TC	CAA	AATTGAAACTCCAGGTACGAG	G	TTGTTTTCTGTGC		243
790	TG	CCTTTAGC	GTACAATCT	C	TTACTTAGAGTAAAATT	CTT	278
EU741688.1	TAC	CCTTTAGC	ATACAATCT	T	TTACTTAGAGTAAAATT	TTT	283
790	ATC	CTTCTCCCGC	CT	A	TTGATGAAGAAGAGAAGCCAC	CTG	318
EU741688.1	GTC	TTTCTCCCGC	TT	G	TTGATGAAGAAGAGAAGCCACT	TTG	323
790	CC						320
EU741688.1	CC						325

Japanese yam mosaic virus isolate FX1, complete genome

Sequence ID: KJ789141.1

JT15	CCATCAACA	GTTGG	CAT	AAC	CAAGTT	CTATGGTGGATGGG	87	
KJ789141.1	GTTGG	TAT	GAC	AAAGTT	TATGGTGGATGGG	7767	
JT15	AT	AGACT	TCTT	GAG	.T..	CCTTACC	AGATGGGTGGA.TA	122
KJ789141.1	AT	GAACT	CCTT	AAT	.AA..	ACTT.CC	GATGGTTGGA.CA	7802

JT15	TATTGTGATGCAGATGGGTCTCAATTTGATAGCTC.CTTG	161
KJ789141.1	CACGTGTGATGCTGATGGCTCTCAATTTGATAGCTC.CTTG	7841
JT15	ACACCTTATCTCATCAATGCGGTGTTGGACATGAG.A.CT	199
KJ789141.1	TCACCATATCTCATAAATGCGGTTTTAAACATACG.A.TT	7879
JT15	AAGTTTT.ATGGAAAGATTGG.GATTTGGGTGAA.CAGA.T	235
KJ789141.1	ACATTTTT.ATGGAGGAGTGG.GATATTGGAGAG.CAGA.T	7915
JT15	GCTGAAA AATTTGTACACAGAAATTGTGTATACAC.CC.A	273
KJ789141.1	GTTGCGTAATTTGTACACAGAAATTGTGTATACAC.CG.A	7953
JT15	TTCTAACNCCTGACGGCACAATTGTCAAGAA GTTCAAANG	313
KJ789141.1	TCGCCACACCGGACGGAACAATAAGTCAAGAAATTCAAAGG	7993
JT15	GAATAACAGTG	324
KJ789141.1	AAATAATAGTG	8004

Lupine mosaic virus isolate LU2, complete genome

Sequence ID: EU847625.2

JT12	AATCAATTCTATGATTTACACACAAAAGGACCATGGAGCG	47
EU847625.2	AATCAATTCTATGATTTACACACAAAAGGACCATGGAGCG	8189
JT12	TTGGAATAACAAAATTTTACGGAGGTTGGAACGAACTTTT	87
EU847625.2	TTGGAATAACAAAATTTTACGGAGGTTGGAACGAACTTTT	8229
JT12	ATCACAACCTGCCAAATGGATGGGTTTATTGTGACGCAGAT	127
EU847625.2	ATCACAACCTGCCAAATGGATGGGTTTATTGTGACGCAGAT	8269
JT12	GGGTCACAGTTTGATAGTTC	167
EU847625.2	GGGTCACAGTTTGATAGTTC	8309
JT12	ATGCAGT	207
EU847625.2	ATGCAGT	8349
JT12	CCCTTTGGGGCATCATATGTTGCGTAAT	247
EU847625.2	CCCTTTGGGGCATCATATGTTGCGTAAC	8389
JT12	ATCATATATACGCCCATATTAGCAGCAGACGGGACAGTTG	287
EU847625.2	ATCATATATACGCCCATATTAGCAGCAGACGGGACAGTTG	8429
JT12	TGAAAAAGTTCCGTGGGAACAACAGTGG	321
EU847625.2	TGAAAAAGTTCCGTGGGAACAACAGTGG	8463

Tobacco rattle virus isolate PeonyTobra1 RNA dependent RNA polymerase gene,
partial cds

Sequence ID: JX144383.1

JT2	CATTCCTGTTGTCCAAATCTAGATATTCGTGATGACCAGA	42
JX144383.1	CATTCCTGTTGTCCAAATCTAGATATTCGTGATGACCAGA	85
JT2	GAGAAATTTCTCGACAAATATTTCTTACTGCTATTGGCGA	82
JX144383.1	GAGAAATTTCTCGACAAATATTTCTTACTGCTATTGGCGA	125
JT2	TCAAGCCAGAAGTGGTAAAAGACAGATGTCGGAGAATGAG	122
JX144383.1	TCAAGCCAGAAGTGGTAAAAGACAGATGTCGGAGAATGAG	165
JT2	CTGTGGATGTACGATCAATTTTCGTGAAAATATTACTGCAC	162
JX144383.1	CTGTGGATGTACGATCAATTTTCGTGAAAATATTACTGCAC	205
JT2	CTAACGCTGTTAGATGCAATAATACATATCAGGGTTGTAC	202
JX144383.1	CTAACGCTGTTAGATGCAATAATACATATCAGGGTTGTAC	245
JT2	ATGTAGAGGGTTTTCTGATGGTAAGAAGAAAGGTGCGCAG	242
JX144383.1	ATGTAGAGGGTTTTCTGATGGTAAGAAGAAAGGTGCGCAG	285
JT2	TACGCAATAGCTCTT.CACAGCCTGTATGACTTCAAGTTG	281
JX144383.1	TACGCAATAGCTCTT.CACAGCCTGTATGACTTCAAGTTG	324
JT2	AAAGACTTGATGGCTACTATGGTTGAGAAGAAACTAGAG	321
JX144383.1	AAAGACTTGATGGCTACTATGGTTGAGAAGAAACTAGAG	364
JT2	TGGTTCACGCTGCTATGCTTTTTGCTCCTGAGAGTATGTT	361
JX144383.1	TGGTTCACGCTGCTATGCTTTTTGCTCCTGAGAGTATGTT	404
JT2	AGTGGACGAAGGTCCATTACCTTCTGTTGATGGTTACTAC	401
JX144383.1	AGTGGACGAAGGTCCATTACCTTCTGTTGATGGTTACTAC	444
JT2	ATGAAGAAGAACGGGAAGATCTATTTTCGGTTTTGAGAAAG	441
JX144383.1	ATGAAGAAGAACGGGAAGATCTATTTTCGGTTTTGAGAAAG	484
JT2	ATCCTTCATTTTCTTACATTCATGACTGGGAAGAGTACAA	481
JX144383.1	ATCCTTCATTTTCTTACATTCATGACTGGGAAGAGTACAA	524
JT2	GAAGTATCTGCTGGGGAAGCCAGTGAGTTACCAAGGGAAT	521
JX144383.1	GAAGTATCTGCTGGGGAAGCCAGTGAGTTACCAAGGGAAT	564
JT2	GTGTTCTACTTTCGAACCGTGGCAAGTGAGAGGAGACACAA	561
JX144383.1	GTGTTCTACTTTCGAACCGTGGCAAGTGAGAGGAGACACAA	604

JT2	TGCTTTTTTCGATCTACAGG.ATAGCTGGAGTTCCGAGGC	600
JX144383.1	TGCTTTTTTCGATCTACAGG.ATAGCTGGAGTTCCGAGGC	643
JT2	GGTCTCTATCATCGCAAGAGTACTACCGAAGAATATATAT	640
JX144383.1	GGTCTCTATCATCGCAAGAGTACTACCGAAGAATATATAT	683
JT2	CAGTAGATGGGAAAACATGGTTGTTGTCCCAATTTTCGAT	680
JX144383.1	CAGTAGATGGGAAAACATGGTTGTTGTCCCAATTTTCGAT	723
JT2	CTGGTTGAATCAACGCGAGAGTNNATCAAGAAANACCTGT	720
JX144383.1	CTGGTTGAATCAACGCGAGAGTTGATCAAGAAAGACCTGT	763
JT2	TTG	723
JX144383.1	TTG	766

Arabis mosaic virus RNA-dependent RNA polymerase mRNA, partial cds

Sequence ID: EU741688.1

787	TTCAGC.ATAAGCACGAG...A.AAGGG.T.AAGCCCGAT	54
EU741688.1GAA...A.AAGGC.A.AAGCCCGAT	60
787	TTTTGTGGGGAGCCTGGAGCGCTGAATTTGTGCCCGG	94
EU741688.1	TCTTTGTAGGGGAGCCTGGAGCAGGGAATTTGTGCCCTGG	100
787	GTGTGGTCCTGAAAAAACTTATTGTAGTCTAGAAGAGGAA	134
EU741688.1	GTGTGGTCCTGAAAAAGCTTATCTGAGTCTAGAAGAGGAA	140
787	TGCAAAACGCGTGTCCCTTCTCTTATAAGTATTGAGACCC	174
EU741688.1	TGCAAAACGCGTGTCCCTTCTCTTATAAGCATTGAAACCC	180
787	CAAAAGATGAAAGGCTTAAAGATCTAAGATTGAAACTCC	214
EU741688.1	CAAAAGATGAAAGGCTCAAGAGGTCCAAATTGAAACTCC	220
787	AGGTACGAGATTGTTTTCTGTGCTGCCTTTAGCGTACAAT	254
EU741688.1	AGGTACGAGGTTGTTTTCTGTGCTACCTTTAGCATACAAT	260
787	CTCTTACTTAGAGTAAAATTCTTATCCTTCTCCCGCTTAT	294
EU741688.1	CTTTTACTTAGAGTAAAATTTTTGTCTTTCTCCCGCTTGT	300
787	TGATGAAGAAGAGAAGCCACCTGCCCAGCTAAGATAGGAA	334
EU741688.1	TGATGAAGAAGAGAAGCCACTTGCCC.....	326
787	TCAA	338
EU741688.1	326

Arabis mosaic virus RNA-dependent RNA polymerase mRNA, partial cds

Sequence ID: EU741688.1

788	TGAAGTCCTTCAGCGAAAAAGCAGGAGA . AAAGGGGTAAAGC	54
EU741688.1	TGAAGTCCTTCAGAGAAAAAGCTGGAGA . AAAAGGCCAAAGC	73
788	CCGATTTTTGTGGGGAGCCTGGAGCGCGTGAATTTGTG	94
EU741688.1	CCGATTCCTTGTAGGGGAGCCTGGAGCACGGGAATTTGTG	113
788	CCCGGGTGTGGTCCTGAAAAAAGCTTATTGAGTCTAGAAG	134
EU741688.1	CCTGGGTGTGGTCCTGAAAAAGCTTATCTGAGTCTAGAAG	153
788	AGGAATGCAAAACGCGTGTCCCTTCTCTTATAAGTATTGA	174
EU741688.1	AGGAATGCAAAACGCGTGTCCCTTCTCTTATAAGCATTGA	193
788	GACCCCAAAAGATGAAAGGCTTAAAGATCTAAGATTGAA	214
EU741688.1	AACCCCAAAAGATGAAAGGCTCAAGAGGTCCAAAATTGAA	233
788	ACTCCAGGTACGAGATTGTTTTCTGTGCTGCCTTTAGCGT	254
EU741688.1	ACTCCAGGTACGAGGTTGTTTTCTGTGCTACCTTTAGCAT	273
788	ACAATCTCTTACTTAGAGTAAAATTCTTATCTTTCTCCCG	294
EU741688.1	ACAATCTTTTACTTAGAGTAAAATTTTGTCTTTCTCCCG	313
788	CTATTGATGAAGAAGAGAAGCCACCTGCCC	325
EU741688.1	CTTGTTGATGAAGAAGAGAAGCCACTTGCCC	344

Liite C: Fylogeneettisten puiden referenssisekvenssit

Taulukko 1. Kuvan 17 fylogeneettisen puun rakentamisessa käytetty nukleotidisekvenssialue kustakin NCBI:n referenssisekvenssistä.

Sekvenssin yksilöllinen tunniste	Nimi	Alue (nt)
EU847625.2	Lupine mosaic virus isolate LU2 complete genome	8150–8463
HG970867.1	Bean yellow mosaic virus gene for polyprotein isolate FB	7569–7882
KJ789141.1	Japanese yam mosaic virus isolate FX1 complete genome	7698–8011
NC_001445.1	Plum pox virus complete genome	7647–7960
NC_001841.1	Sweet potato feathery mottle virus complete genome	8713–9026
NC_002509.2	Turnip mosaic virus complete genome	7832–8145
NC_002634.1	Soybean mosaic virus complete genome	7608–7921
NC_003224.1	Zucchini yellow mosaic virus complete genome	7615–7928
NC_003398.1	Sugarcane mosaic virus complete genome	7464–7777
NC_003537.1	Dasheen mosaic virus complete genome	7875–8188
NC_003605.1	Lettuce mosaic virus complete genome	8099–8412
NC_004013.1	Cowpea aphid-borne mosaic virus complete genome	7483–7796
NC_004047.1	Bean common mosaic necrosis virus complete genome	7658–7971
NC_004752.1	Yam mosaic virus complete genome	7608–7921
NC_006262.1	Watermelon mosaic virus complete genome	8014–8327
NC_006941.1	Cucumber vein yellowing virus complete genome	7541–7854

Taulukko 2. Kuvan 18 fylogeneettisen puun rakentamisessa käytetty nukleotidisekvenssialue kustakin NCBI:n referenssisekvenssistä.

Sekvenssin yksilöllinen tunniste	Nimi	Alue (nt)
X06172.1	Tobacco rattle virus RNA-1 complete sequence	614–1334
NC_003805.1	Tobacco rattle virus RNA 1 complete sequence	614–1334
AF406990.1	Tobacco rattle virus segment RNA1 complete sequence	614–1334
AF314165.1	Tobacco rattle virus strain ppK20 segment RNA1 complete sequence	614–1334
AF034622.1	Tobacco rattle virus RNA polymerase 29 kDa protein and 16 kDa protein genes complete cds	613–1333
JX144383.1	Tobacco rattle virus isolate PeonyTobra1 RNA dependent RNA polymerase gene partial cds	46–766
GQ903771.1	Tobacco rattle virus isolate MI-1 segment RNA1 complete sequence	614–1334
KJ826365.1	Tobacco rattle virus strain HaB segment RNA 1 complete sequence	591–1311
KF758793.1	Tobacco rattle virus isolate Slu24 segment RNA1 complete sequence	614–1334
HM195288.1	Tobacco rattle virus replication-associated protein 2 replication-associated protein 1 movement protein and 16K protein genes complete cds	591–1311
X14006.1	Pea early browning virus (PEBV) genomic RNA1 for 141K 201K 30K and 12K proteins	536–1277

Taulukko 3. Kuvan 19 fylogeneettisen puun rakentamisessa käytetty nukleotidisekvenssialue kustakin NCBI:n referenssisekvenssistä.

Sekvenssin yksilöllinen tunniste	Nimi	Alue (nt)
NC_006057.1	Arabidopsis mosaic virus RNA 1 complete sequence	5024–5338
NC_003509.1	Blackcurrant reversion virus RNA1 complete sequence	4638–4961
KC138732.1	Arabidopsis mosaic virus isolate NW segment RNA1 polyprotein 1 gene complete cds	5021–5335
JQ975057.1	Arabidopsis mosaic virus isolate DU13 segment RNA1 complete sequence	5027–5341
JF968120.1	Grapevine fanleaf virus isolate SAPCS3 segment RNA1 complete sequence	5043–5357
GQ369528.1	Arabidopsis mosaic virus isolate ba segment RNA1-III complete sequence	5031–5345
GQ369527.1	Arabidopsis mosaic virus isolate ba segment RNA1-II complete sequence	5032–5346
GQ369526.1	Arabidopsis mosaic virus isolate ba segment RNA1-I complete sequence	5031–5345
EU741688.1	Arabidopsis mosaic virus RNA-dependent RNA polymerase mRNA partial cds	14 –328
EU617326.1	Arabidopsis mosaic virus isolate Lv segment RNA1 complete sequence	5032–5346